THE PSYCHOLOGICAL BASIS OF AXIOMATIC MATHEMATICS

By

Floyd Grant Robinson

1959

UNIVERSITY OF ALBERTA
LIBRARY

Regulations Regarding Theses and Dissertations

Typescript copies of theses and dissertations for Master's and Doctor's degrees deposited in the University of Alberta Library, as the Official Copy of the Faculty of Graduate Studies, may be consulted in the Reference Reading Room only.

A second copy is on deposit in the Department under whose supervision the work was done. Some Departments are willing to loan their copy to libraries, through the interlibrary loan service of the University of Alberta Library.

These theses and dissertations are to be used only with due regard to priority rights of the author. Written permission of the author and of the Department must be obtained through the University of Alberta Library when extended passages are used. When permission has been granted, credit must appear in the published work.

This thesis or dissertation has been used in accordance with the above regulations by the persons listed below. The borrowing library is obligated to secure the signature of each user.

Please sign below:

Date	Signature	Institution

THE UNIVERSITY OF ALBERTA

THE PSYCHOLOGICAL BASIS OF AXIOMATIC MATHEMATICS

A DISSERTATION

SUBMITTED TO THE FACULTY OF GRADUATE STUDIES

IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE DEGREE

OF DOCTOR OF PHILOSOPHY

DIVISION OF EDUCATIONAL PSYCHOLOGY

by

FLOYD GRANT ROBINSON

EDMONTON, ALBERTA

APRIL, 1959

UNIVERSITY OF ALBERTA

FACULTY OF GRADUATE STUDIES

The undersigned hereby certify that they have read and recommend to the Faculty of Graduate Studies for acceptance, a thesis entitled, "The Psychological Basis of Axiomatic Mathematics," submitted by Floyd Grant Robinson, M.A., B.Ed., in partial fulfilment of the requirements for the degree of Doctor of Philosophy.

PROFESSOR

PROFESSOR

PROFESSOR

PROFESSOR

PROFESSOR

PROFESSOR

PROFESSOR

April, 1959.

ABSTRACT

This thesis arose from an attempt to resolve a pedagogical dilemma faced by the secondary school mathematics teacher by providing a common ground for the research-oriented psychological approach of the elementary mathematical education, and the pure mathematics orientation of advanced mathematical education. It seemed that a particularly useful type of theory would give emphasis to the limits which commonly appear in the student's mathematical performance.

To meet this need, a theory of mental operations was proposed which employed the notion of an underlying hypercathexis energy. This energy, which results from the conservation of primitive cathexis through increased mental organization, is used to power the independent thinking of the super-organization of central processes. The possession of hypercathexis allows the organism to deal in an anticipatory way with its environment through the activation of neural structures which represent an internalization of systems in the exterior world.

Two principal schema of mental operation were postulated. In Type A functioning, hypercathexis is expended in a sequence of 'operation-storage' couplets. Progress in this sequence would seem to be potentially unlimited,

although several causes of induced upper bounds were foreseen. In Type B functioning, which is encountered in situations involving logical or utility choice, the available hypercathexis is divided between alternatives, with the result that natural upper bounds can be expected.

The structure of axiomatic mathematics was next considered and mathematical ability was defined in terms of the individual's expected performance with respect to this structure. The ability to comprehend mathematics was denoted Level I ability and was found to involve Type A mental functioning. A second type of ability (Level II) was required to deal with gaps in the deductive sequence. If the individual filled the gap by a non-random procedure which generated successive choice points, then his behaviour was described as problem-solving (Level II); this ability was seen to involve Type B functioning. A treatment of Level III ability, the creative aspect of mathematical performance, was considered to be outside the range of our available concepts.

The experimental work was designed to corroborate various aspects of the theory including several predictions made from it. A semi-clinical approach was employed in which a battery of standardized and experimental tests was administered to 106 high school students.

In three studies at Level I, the comprehension time was found to be a non-linear function of demonstration length. Two 'factors' emerged in this area; the most important one was closely related to the speed of performing simple mental operations and exhibited high correlations with Otis IQ. This was interpreted to mean that the speed of performing simple mental operations was dependent upon hypercathexis level. A further experiment, which employed comparisons of grade 10 and grade 12 groups matched in Otis IQ, suggested that the mental energy level does not increase once the stage of formal mental operations is reached, and that further gains in mental performance may be attributed to the development of energy-conserving strategies. The experiments also supported the theory in showing that group pressures may cause a lowering of upper bounds in Level I performance.

The studies at Level II were largely devoted to the solution of geometry problems. The experimental results again corroborated the theory in that 90% of the students demonstrated upper bounds which were predictable within the experimental time limit. Again, the student who made continued progress did so by categorizing the propositions in such a manner that a continuous refinement of strategy balanced the increase in the number of propositions. Once the student

reached an upper bound, he tended to add propositions to his repertoire without increasing his power of analysis. A further study of logical behaviour at and above the upper bounds gave rise to serious doubts concerning the alleged value of geometry as a training in the techniques of correct thinking. It also appeared that a priori prediction of geometry problem-solving ability was not feasible, but that a 'trial' method would likely be useful.

A further investigation of the overachiever in geometry was undertaken. This study revealed that many students of mediocre IQ tended to do well in geometry and that the key to the overachievers' success lay in an intense motivation which was manifested in a willingness to persevere. There was some indication that the source of this extreme ego-involvement with mathematics could be traced to a period of social maladjustment.

ACKNOWLEDGMENTS

The author wishes to convey his gratitude to the many people who have contributed in one way or another toward making this thesis possible. He is particularly indebted to the supervisory committee, to it's chairman Dr. C. C. Anderson, and to the divisional chairman Dr. G. M. Dunlop for their interest in the thesis and for the encouragement they have given over the past two years.

TABLE OF CONTENTS

LIST OF TABLES

TABLE PAGE

LIST OF FIGURES

CHAPTER I

INTRODUCTION

The contemporary high school teacher of mathematics faces mamy unsolved pedagogical problems. His domain--high school mathematics--has remained practically invulnerable to many attempts to uncover its psychological dimensions. This thesis attempts to introduce a unified point of view in this area, by providing a theory of mathematical ability which unites current psychological principles with mathematical content.

In order to appreciate the necessity of the type of theory proposed, and to allow a comprehensible statement of the thesis problem, it will be necessary to make a brief preliminary survey of certain critical influences which bear upon the high school mathematics teacher. We consider these forces as emanating from two principal sources: (a) elementary mathematical education, and (b) advanced mathematical education. Some characteristics and lines of influence of each will be considered in turn.

Elementary Mathematical Education

The most outstanding development in this area in the last fifty years has been the research-reorientation of educational theory and practice. This has made possible a consolidation of opinion on certain basic issues.

1. Content and Grade Placement. There has always been agreement that the elementary school curriculum ought to teach the number facts and skills necessary to carry on the computations of everyday life. Educational research has buttressed this position by determining the content which is appropriate to a particular age level. Readiness studies (Buckingham, 1930; Gunderson, 1940) have ascertained the child's ability to handle number concepts on entering school, and have made possible a systematic arithmetic program in grades one and two. Again, the proper grade placement of topics has been extensively investigated (Smith, 1950; Thiele, 1938), so that while agreement is not complete (Washburne, 1930; Washburne, 1939), there is a sufficient body of experimental evidence to defend current practices.

2. Learning Theory. The first systematic psychology of arithmetic based on a well-developed theory of learning was put forward by Thorndike (1922). This rigid connectionism (based on the law of practice) was, in time, attacked by exponents of the Gestalt position (Wheeler, 1935) so that the years 1930 to 1950 witnessed an extremely heated and outspoken contest in the professional literature (Brownell, 1935; Knight, 1930; McConnell, 1941; Skinner, 1954). Nevertheless, by 1952, there was fairly general agreement that arithmetic should be taught "meaningfully", i.e., in accordance with the meaning theory put forward by Brownell (1935).

The 'meaning' theory conceives of arithmetic as a closely knit system of understandable ideas, principles, and processes. According to this theory, the test of learning is not mere mechanical facility in 'figuring'. The true test is an intelligent grasp upon number relations and the ability to deal with arithmetical situations with proper comprehension of their mathematical as well as their practical significance.

3. Specific Classroom Practices. A large proportion of the 1,413 research papers published by the end of the year 1945 (Buswell, 1951) have substantiated experimentally, specific procedures for the classroom. Representative of the type of research done is Brownell and Mosers' well-known proof of the superiority of the 'decomposition' method of subtraction over the 'equal-additions' method.

Influence on Secondary Education

No one will claim that the field of elementary mathematical education has solved all its problems. Yet it has accumulated a sufficient body of proven facts and principles to offer its teachers a framework within which they may work with confidence. The high school teacher of mathematics, in so far as his training includes educational psychology, is exposed to principles which have--as far as mathematics is concerned--been tested and found applicable only at the elementary level. For example, when educational psychology textbooks discuss basic principles of learning theory, they usually fail to distinguish between applications at the

elementary and secondary levels, despite the fact that few applications have been made beyond the early stages of the elementary grades (Dyer, 1956). Consequently, it may be said that one disturbing force to which the high school teacher is subjected, is the expectation that he will--or can--apply in his area, the results or principles derived from research conducted in another area. Indeed, it turns out that research at the secondary level has been noticeably unsuccessful, and many high school teachers have concluded (Savage, 1956) that research findings are not applicable to their work at all.

Advanced Mathematical Education

In sharp contrast to the psychological orientation at the elementary level, mathematical education at the college and university proceeds almost entirely without reference to psychological principles in questions of content and placement, learning theory, and specific lecturing procedures. Occasional pseudo-psychological excusions into mathematical creativity (Hadamard, 1945), multiple-correlation predictions of college marks from high school mathematics grades (Douglass, 1936), and a few factor-analytic investigations of mathematical ability at the college level (Coleman, 1956), fairly well represent the extent of the influence of psychological thinking at the advanced level. Indeed, it is likely that most mathematicians regard their

work to be philosophically rather than psychologically oriented.

1. <u>Control of Curriculum</u>. The mathematician's main line of influence in secondary education has come through control of the curriculum--a curriculum which has been considered by the mathematician as preparatory to university mathematics. Historically, the last major revision in the high school mathematics programme in the United States, dates back to the Committee of Ten (1894). In this respect, Kinney and Purdy (1952) remark,

> Because the purposes of the sequence were cultural and disciplinary, primarily designed for college preparation, it is not surprising to find that no psychologist or sociologist was included on the committee, and that none of the committee members had a record of close contact with youth. This may account for the fact that no differentiation in purposes was recognized between the secondary school and college.

According to a recent study (Dyer, 1956), the only major change since that date has been the gradual disappearance of solid geometry, and the introduction of a general mathematics course at the grade 9 level.

The mathematician's influence over the high school curriculum has been further strengthened by the control of standards (through departmental examinations), the writing of textbooks, and the control of courses for the mathematical training of prospective teachers.

2. The Effect of Increased Enrolment. At the university level, increased enrolment probably means that more students are being provided with an opportunity to continue their education; it does not necessarily imply, however, that the quality of the student entering is declining. At the high school level, on the other hand, the average I.Q. of the incoming student has most certainly declined over the past few decades, since the percentage of students entering high school at the present time is more than eight times as large as it was in 1900 (Kinney, 1952, p.33). Thus, mathematics courses designed for approximately 10% from the upper part of the intelligence distribution are now offered to a group which includes a large proportion of students of mediocre ability. This condition has only partly been offset by the offering of more options in the larger schools. Under the circumstances, it is not surprising that many students are not able to cope with the mathematics courses offered to them.

In summary, it may be said that a second disturbing force acting on the contemporary mathematics teacher arises from the fact that he is expected to teach university preparatory courses designed for an intelligent minority, to a relatively unintelligent majority, most of whom will never progress through, or beyond, the high school.

Secondary Mathematical Education

The unsatisfactory position of the high school mathematics teacher has already been indicated--and will become clearer by a consideration of the three areas listed under elementary education.

1. Content and Placement. The early studies of Thorndike (1924) discredited the then-current belief in the 'mental discipline' values of algebra and geometry. Since that time, curriculum planners have been careful to exclude from the proposed aims of mathematics courses any statement which might be construed as a belief in 'faculty' psychology. However, it is still clear that a great deal of transfer is expected. For example, one of the six objectives for the geometry course as set down in 1940 by the Joint Commission to Study the Place of Mathematics in Secondary Education, was the following: "Understanding of deductive science and ability to apply the deductive method to non-geometric arguments". After reviewing studies (Ulmer, 1939) which attempt to verify such claims, Dodes (1953) points out that the 'home-made' tests used to verify transfer are of doubtful validity, and consequently the supposed merits of the high school mathematics curriculum as a promoter of desirable mental attitudes or abilities remain in the realm of supposition rather than experimental fact.

Not only has the psychological value of the curriculum been cast into doubt; its mathematical value is now maintained to be questionable. At the present time, there is a movement in progress to "modernize" the high school mathematics curriculum. The current literature abounds with articles by professional mathematicians (Van Duren, 1958) who point out that the ideas with which the secondary school mathematics student has contact are of ancient vintage--predating in fact, Newton's invention of the calculus. Moreover, they contend that large sections of the algebra and geometry courses could be deleted with no ill effects--even for those few proceeding to university mathematics. The time saved would be devoted to such modern ideas as set theory, symbolic logic, and statistics. Although similar themes have been heard before, (Thorndike, 1923), it is possible that the next few years will witness some major revisions in the high school curriculum.

2. Learning Theory. It has already been pointed out that learning theory has been little applied to mathematics, beyond the elementary level. Thorndike's (1923) discussion of algebra from the connectionist point of view, and Wertheimer's (1945) applications of Gestalt psychology to geometry represent the main applications of pure learning theory at the secondary level. It is, therefore, probably

not unfair to say that the high school mathematics teacher makes little use of any systematic theory of learning in his teaching (Dyer, 1956).

3. Specific Practices. Most attempts to justify the specific practices of the high school classroom have yielded negative or inconclusive results. The results of experiments to determine: the optimum size of classes (Schunert, 1951), the superiority of deductive or inductive methods of teaching (Michael, 1949), the optimum amount of homework and the effect of easy and hard marking (Schunert, op. cit.), the optimum frequency of tests (Johnson, 1938), and the effect of the academic training, background, and experience of the teacher on the achievement of the pupils (Schunert, op. cit.), have not endorsed or supported any particular practice to the exclusion of others.

The full implication of the position of the high school mathematics teacher is now clear. For while his mathematical training and subject matter field are oriented toward the axiomatic systems which are characteristic of modern mathematics, his pedagogical training is oriented toward psychological principles which have proved successful only in the elementary area and which do not seem to apply to secondary school mathematics. Unfortunately, the lack of congruence which exists between these bi-polar

forces has left the high school teacher in a state of uncertainty, without a frame of reference with which to defend or examine his classroom practices.

One is tempted to say that a solution to these problems will eventually come through an improvement in learning theory, and one expert has declared that we must rebuild our learning theories, moving beyond the 'well-controlled-animal' point of view to a 'complex-human-being' theory (Buswell, 1956).

There are many aspects of learning theory (Hilgard, 1957, p.8); the emphasis given to the various parts undergoes a change as we move up the educational scale. In the elementary school, where we assume that it is necessary for every child to learn a certain minimum amount, the question of _how_ the individual learns has been given the greatest attention, but in the selective practices of the high school and university, the emphasis shifts to a consideration of _how much_ the individual can learn. The latter question is certainly of crucial importance to the high school mathematics teacher who has to deal with large numbers of students who seem to be unable to progress beyond the very elementary stages in mathematics. A theory of mathematical ability which would allow the prediction of an individual's progress to a certain level in an axiomatic system is badly needed at the present time. Such a theory,

in providing a framework and concepts for the direction of research in secondary school mathematical education, would go a long way toward a solution of the outstanding problems in that area.

Statement of Thesis Problem

This thesis attempts to advance and substantiate a theory of mathematical ability with the following characteristics:

(a) It should be capable of alleviating the bi-polar stress on the mathematics teacher by describing performance in mathematics (particularly in the axiomatic systems of modern mathematics) in terms of current psychological principles.

(b) It should explain the apparent limits in the mathematical performance of students in terms of cognitive, motivational, and classroom variables.

(c) It should distinguish between the types of mental tasks required in mathematics and should describe an individual's limit in terms of attainment of a certain level in a scale of difficulty in each type of task.

Section A of the thesis is devoted to the development of this theory, and Sections B and C to its experimental verification.

SECTION A

A THEORETICAL CONSIDERATION OF THE PSYCHOLOGICAL BASIS OF AXIOMATIC MATHEMATICS

CHAPTER II

A THEORY OF MENTAL OPERATIONS

This chapter presents those ideas concerning the operation of the brain which are prerequisites to the development of the theory in the following chapters.

Use of Models

Most of the theorizing in this thesis--and in contemporary psychology--may be viewed as degrees of approximation to 'model construction'. A model[1] represents a way of dealing with complex phenomena by means of a comparison with an analogous situation, usually one in mathematics or physics. The process may be schematized as in Fig. 1.

After simplifying some 'real-world' situation S by reducing it to its essential variables, the model constructor then maps S in a one-to-one fashion on a mathematical or physical system $\bar{S}$. Once the mapping is complete, he can apply to $\bar{S}$ established theorems of mathematics and physics, proceeding without obligation to interpret these operations in terms of 'events' in the real world. When a result $\bar{R}$ is

---

[1] A more thorough account of the logical relationship between theory and model may be found in Braithwaite (1953, Ch. IV).

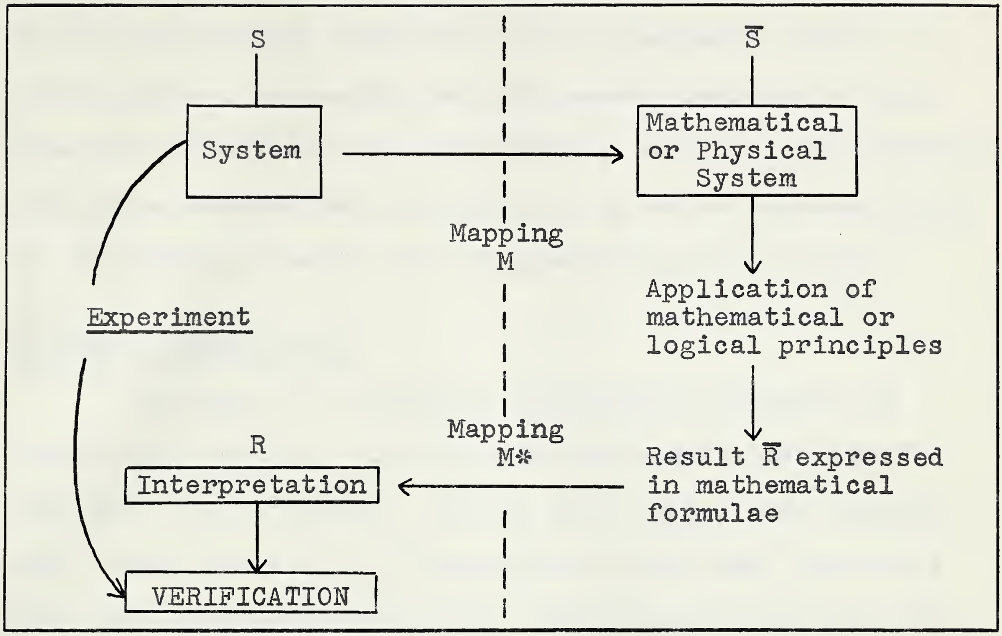


Fig. 1. Schematic representation of model construction.

obtained, it is mapped back onto R and is then subjected to interpretation: if it yields results not given as properties of the original system S which are potentially verifiable by means of an experiment performed on S, then the model is a <u>useful</u> one. If the model yields results which are not new, or results which contradict such experimental evidence, then the model is a <u>useless</u> one, and may be discarded.

Although model construction would seem to be inevitable in scientific thinking, most of our psychological models lack explicit mathematical formulation. Consequently,

it is often not clear when we use a term, whether we imply by it some property intrinsic to the phenomena under investigation, or whether we have merely borrowed a term from the nomenclature of the physical analogy. Nevertheless, with these reservations in mind, we go on to construct what is, in effect, a model for the operation of the brain.

Levels of Description

The gamut of scientific description advances by increasing levels of organization and complexity from the subatomic to the cosmic. Within this range, the scientist must choose the level of description which most naturally fits the phenomenon which he is investigating at any given time. In particular, the psychologist has found it advantageous to describe human behaviour by employing neural, psychic, and social models.

Proceeding up the scale of increasing complexity of organization, we encounter at each level, phenomena which cannot be completely analysed into the units of the level below. Thus the ego construct at the psychic level is not equatable to any known compound of neural structures. This state of affairs is not surprising, for if complete decomposition were possible, then a single system of concepts would suffice in psychology. As things stand, however, theorists have developed separate systems of nomenclature

and constructs appropriate to each level. Moreover, the natural desire for 'complete' systems (Hilbert & Ackermann, 1950) has not only resulted in the development of single-level systems (i.e., employing concepts at one level), but has also created a taboo against multiple-level theorizing. Thus a psychologist using physiological concepts tends either completely to discredit psychic concepts, or to reduce them to physiological terms.

In many cases, it may be possible for the pure psychologist to operate successfully with a one-level approach. This is particularly true in laboratory situations which have the effect of reducing the dimensions of behaviour by controlling conditions so that psychic and social variables can be ignored. However, the controlled laboratory method has been disappointing in its contribution to educational psychology, so disappointing, in fact, that Buswell (1956), in speaking of learning theory, implied that this type of investigation may be invalid as a technique for studying human behaviour at the level of complexity which is necessary for educational practice.

The problem investigated in this thesis illustrates the inadequacy of a one-description-level point of view. We are interested in the problem solving behaviour of the student within the mathematics classroom. In particular, we are concerned with the extent to which performance is

bounded. One major determinant of upper bounds is the physiological structure of the brain together with the operations which it is able to perform, because brain structure must necessarily set limits to the work which the brain may perform. Actual performance, however, depends as well upon such motivational variables as ego-involvement, persistence, and level of aspiration. Thus, psychic concepts must be considered. Moreover, since the student is influenced by others in the group, we are compelled to introduce interaction concepts as well.

When educational psychology achieves the status of a more exact science, one of the problems with which it will first have to deal is the extent to which concepts from the three levels may be used in the description and explanation of a single piece of behaviour. At the present time, we use the 'layer' method of description. Thus standard works in educational psychology discuss physical, mental, and social aspects of behaviour separately, confining the separate descriptions to different sections of the text. A complete description will require an integrated treatment, which will show how neural, psychic, and social phenomena interest to determine performance at a specific time.

Energy

The notion of energy cuts across the boundaries of the three levels of description, and for this reason would seem to be the natural concept to employ in describing interaction. The problem then reduces to a consideration of the relationship between the electro-chemical energy of the physiologist and the psychic energy (cathexis) of such neo-Freudian theorists as Rapaport (1951), and Colby (1955).

Freud originally conceived of cathexis in chemical and electrical terms, and did not modify his position radically in his later thinking. His final remarks on energy appear in his Outline of Psychoanalysis (1940):

> We assume, as other natural sciences have taught us to expect, that in mental life some kind of energy is at work; but we have no data which enable us to come nearer to a knowledge of it by analogy with other forms of energy.

More recently, psychoanalysts have become concerned about this concept, and there has occurred a tendency to present cathexis as a postulated form of energy "in addition to known mechanical, thermal, electrical, and chemical forms" (Colby, 1955, p.28). He further maintains that "cathexis energy cannot as yet be converted into other forms of energy by any logical or empirical means". He adds, however, "But maintaining the continuity of integrative levels in science, we can say that cathexis energy (CE) bears some relationship to other forms. For example, CE at

least does not disobey the principles formulated for other forms of energy".

The central concept of 'drive' offers further illumination concerning the relations which may be postulated between cathexis and 'ordinary' energy. Colby (ibid., p. 38) describes drives as originating in certain internal biochemical processes which are "somehow brought into connection with drive schemas", i.e., structural components of the psychic apparatus which are powered by cathexis. But since the biochemical processes are themselves describeable in ordinary energy terms, the arousal of cathexis in this psychic system is correlated with a physiological energy change. Again, when the drive schema is energized, it impels the "motion or action of the organism toward the environment" (ibid., p. 38). Here, an expenditure of cathexis in the psychic system is correlated with an expenditure of physiological energy in the neural system.

Thus it often happens that, when cathexis energy transformations are postulated within the P.A. at the psychic level of conceptualization, ordinary energy transformations occur in the physiological system at the neural level of conceptualization. When, in describing behavioural phenomena at the neural level we invoke the particular form of physiological energy which Hebb (1949)

denotes 'central facilitation', we can, and will assume that if we were to shift our frame of reference to the psychic level, we should often find cathexis energy in play.

The problem of a suitable language in which to describe a dual system of constructs is a very perplexing one and would probably be best handled by developing a meta-system in which the object of study becomes the language and permissible combinations of constructs of the various levels of description.[2]

No such formalization has been attempted here. However, when a statement such as "phase sequences may be activated by a cathexis powered central facilitation" occurs, it is to be understood as a statement in the language of the meta-system which is intended to convey the idea that the activation of a phase sequence (a neural construct) can be accomplished by a form of energy which is called central facilitation at the neural level of conceptualization, but which implies the expenditure of cathexis in a psychic conceptualization of the same process.

---

[2]This would present a field similar to meta-mathematics as outlined by Kleene (1952).

The Homeostasis Model

An organism is conceptualized as a type of generalized machine in constant interaction with its environment. The machine consists of a set of interrelated variables (body temperature, pulse rate, etc.); both the organism and its environment may increase in complexity during the life history of the organism through a gradual extension of the organism's range of sensitivity to include 'interaction' or 'social' forces in the environment.

The fact that the organism maintains a rather remarkable stability in the face of an extremely changeable environment has given rise to the concept of homeostasis (Cannon, 1920). According to this principle, the displacement of a variable from its position of equilibrium is conceptualized as a 'need'--and is counteracted by an act of the organism which brings into play additional variables--the overall effect being to confine each variable to a region which lies within certain 'critical limits'[3]. The organism has at its disposal a quantity of primitive energy--usually called psychic energy (Colby, op. cit.,

---

[3]Ashby (1952) demonstrated how a single variable may be controlled by the enlargement of the variable phase-space through the introduction of a second variable. He conceived of the brain as performing a rapid sequence of switching operations until the phase point is located within a region of stable equilibrium.

p. 28)--which it utilizes to maintain its overall equilibrium. The displacement of a variable beyond a threshold level determined by the complete configuration of variables at the given time, results in a state of drive, and in the utilization by the organism of its available energy to reduce this drive[4].

Evolution of the Brain

The history of the evolution of the brain is a history of the growth of independence of the organism from the environment. Elementary organisms are environment bound, i.e., obligated to expend energy to obtain incentives (need satisfiers) in their immediate environment. Any radical change in the environment may result in uncontrollable displacements in some essential variable--and the extinction of the organism.

The independence of the higher organism is reflected in two ways:

(a) The higher organism is able to explore the present environment without physical involvement. The evolution of differentiated sense organs allows the organism to explore environmental conditions without exposing its main body to unknown and potentially dangerous conditions of the

---

[4]Freudian theory refers to the investment of energy in instincts as 'cathexis' (Hall and Lindsey, 1957, p. 36).

environment. With the later development of the distance receptors, the environment could be explored up to any distance within the range of these receptors[5].

(b) The higher organism is able to envisage the results of its future actions in the environment by enacting them internally in 'miniature' and thereby is able to choose appropriate actions with regard to their consequences in terms of the future need-structure of the organism. This anticipation of future needs relieves the organism from the constant necessity of immediate adjustment. By this means, the immediate adjustment of the primitive (and human infant) organism is extended to the long range adjustment of the mature human organism[6].

Maturation of the Brain

Since the human infant at first resembles the primitive organism, there remains to be explained: 1. how the mind gradually acquires the ability to perform function (b), and 2. the sort of mental functions or operations involved

---

[5]This position is taken by Thurstone (1927, p. 69) who considers the phylogenetic development of intelligence in the light of the principle of 'protective adjustment'.

[6]According to Rignano (1923, p. 19),
"By increasing the power of predicting external events, the intellect succeeds constantly in devising new means, more indirect and more complex for attaining certain ends."
Similar themes are expounded by Binet (1904) and Piaget (1951).

in this performance. From birth onward, there is a gradual mapping of the exterior environment on the sensory cortex of the brain[7]. Since the association areas and neural pathways of the brain are finite and the environment with which the organism will eventually have to deal is potentially infinite, the mapping must proceed in a many-to-one fashion. This type of mapping would result in the formation of equivalence classes, i.e., sections of the environment which the organism can, in some respect, treat as equivalent. Primitive equivalence classes are formed around the basic needs (e.g. class of edible foods, class of dangerous objects), but the process later seems to become autonomous. The physiological process involved in the formation of equivalence classes is difficult to describe briefly, but it involves the establishment of related groups of neural circuits (phase sequences) by virtue of their common central

---

[7]Hebb's (1949) theory, based on contemporary knowledge of the physiology of the brain, offers a plausible account of how continued stimulation from the environment causes a <u>non-random</u> lowering of synaptic resistance (<u>neurophysiological</u> postulate, p. 62). Thus the superposition of external stimulation upon the spontaneous firing of the association areas (p. 121) results in the establishment of primary electrical circuits--the so-called 'cell assemblies'.

facilitation through the cathexis (i.e., primitive energy) accruing to a particular need[8].

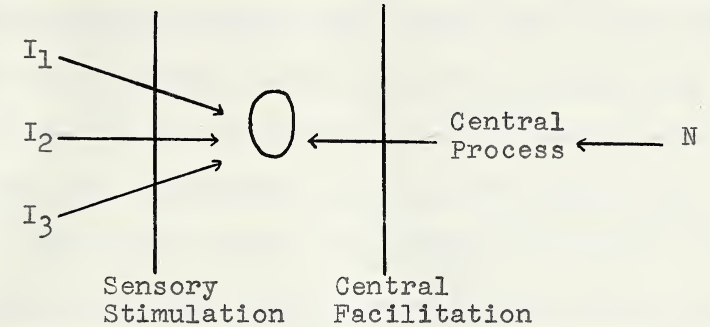


Fig. 2. Energy sources in the formation of phase sequences.

Each phase sequence is formed under the influence of two sources of stimulation: the sensory stimulus from the i'th incentive I_i, and a central facilitation C, whose source of energy we assume to be the cathexis accruing to

---

[8]Rapaport's (1951) theory, which emphasizes the central aspects of the organization of the organism has difficulty explaining the secondary process--the contact of the organism with reality. Hebb's theory explains this domination of the brain by the environment, but is less successful with (and minimizes) the influence of the primitive structure and energies of the organism. However, since the formation of phase sequences is influenced by both sources, (Hebb, 1951, p. 121), it would seem feasible to combine both points of view by supposing that the hypercathectic energies of Rapaport are manifested in the spontaneous neural firing which provides the basis of central facilitation.

the common need N. Since a given need could presumably manifest itself--as central facilitation--in a particular pattern of spontaneous firing, we might expect more than a chance relationship between the neural loci of the phase sequences corresponding to the incentive I_1, I_2, and I_3.

Objects in the environment can, of course, belong to more than one equivalence class. Not only do the classes intersect, but there exists a hierarchy of classes of increasing generality and inclusiveness[9]. As soon as a rudimentary system of classes has been established, the organism tends to classify its subsequent experiences according to the existing categories[10].

This type of organization of behaviour implies a saving of the cathectic energies powering the need system, for as the members of a class are functionally equivalent, they may stand one for another in relation to the satisfaction of a need, thereby eliminating the necessity of a

---

[9]This argument is in harmony with Bruner (1956, 1957) who treats categorization as the principle cognitive function of mental operation.

[10]If two items of experience are compared, the most economical method of comprehending them would be to relegate to them a class common to both. However, since a class is defined by a relation common to all its elements, there would seem to be a tendency, based on the natural mode of operation of the brain toward the establishment of relations. This closely approximates Spearman's (1927) principle of the 'education of relations'.

search for a specific member of the class. The surplus energy, released to a super-organization of the central processes--usually conceptualized as the ego (Rapaport, p. 702)--is used to carry on the 'reality testing' function (b) of the 'secondary process'. The super-organization of central processes (S.C.P.) uses part of its energy to gain freedom from the immediate and disruptive demands of the organism's needs by a process which, in effect, neutralizes part of the energy invested in that need--thereby keeping its potential below the level required for entry into consciousness.

The remaining energy (attention cathexis)--a quantity which would normally fluctuate depending on the momentary overall displacement of the need variables--can be used for the independent (i.e., independent of need) operation of the S.C.P. The S.C.P. also gains partial independence from external stimuli, for while these stimuli are necessary in the formation and consolidation stages of phase sequences, central facilitation by itself may suffice to activate the mature phase sequence. Consider, for example, the physical act of moving a small pebble from the left edge of a sheet of paper to the right edge. Through the performance of similar or near-similar acts many times in the past, the operation consisting of perception of object, physical movement, and perception of result, is organized into an

integrated pattern of phase sequences.

When the act is performed mentally[11], a phase sequence is activated which belongs to what might be termed a representative of the class of pebbles. When attention-cathexis is first directed to the neural locus corresponding to the concept 'pebble'--there seems to be a tendency for central facilitation, to activate the whole neural locus; however, continued 'attention', i.e., central facilitation, results in the activation of a specific representative or locus[12].

---

[11] It will be convenient to use the adverb 'mentally' to refer to centrally-directed operations.

[12] If the neural locus representing a concept consists of individual circuits whose potential decays with time, then after the initial activation of the whole field, the activity of individual circuits would tend to decline at a rate which would vary directly as the circuit resistances. However, the resistances tend to be broken down through repeated activation--and the most frequently encountered member of the class would have a phase sequence whose potential decays more slowly than others. Thus it would follow that the first global impression of a concept soon particularizes into a specific representative. This would seem to be true of concepts (house, fender) whose specific instances are clearly distinguishable. For concepts whose specific instances are not distinguishable (line, point), the representative probably stands for a 'neural average' of cumulative experience. There would, moreover, be required a constant expenditure of retention-cathexis to maintain the concept in a 'general' form. This point of view was advanced by Thurstone (1927) in his theory of 'inhibition'.

The potential of the centrally activated phase sequence corresponding to the representative[13] would be much lower than the phase sequence resulting from the perception of a pebble in the exterior environment, since the potential of the latter phase sequence would be increased by a sensory stimulation.

Similarly, the phase sequences corresponding to the imagined physical act are activated in their turn, at a potential sufficient to activate the final phase sequence--viz., the visualization of the customary[14] result, but insufficient to stimulate the musculature to overt physical action[15]. Thus the phase sequences corresponding to the

---

[13]'Imagining' need not always be carried on in terms of representatives. It may be possible, for example, to imagine any particular non-representative stone. But this imagining would necessitate a larger expenditure of energy, and would perhaps even require direct support from sensory stimulation. Thus we can imagine a chair, moving across a room, but to imagine this particular chair may require looking at it first.

[14]Rignano (1923, pp. 170-171) dealt with the necessity of a result following its antecedent action:

> That once the tendency of a given quantity towards a given limit is recognized, we can imagine this last as already attained by the quantity itself, and take the limit simply in place of this latter--this rests upon the general psychological fact for all reasoning whatever, that the fact of seeing in imagination the possibility of performing certain operations or experiments leads naturally to their being imagined as already performed.

[15]This point of view is supported by the work of Jacobsen (1932) and Max (1935) who demonstrated that the process of thinking or imagining is accompanied by the arousal of minute electrical currents in the musculature concerned.

mental act are low-potential repetitions of the phase sequences usually activated in the corresponding physical act.[16]

It follows that the operations which the mind may imagine as performed on a representative are those resulting from possible physical operations performed on a class member.[17] This fact will be important in a subsequent study of the operations in axiomatic mathematics.

---

[16]Some comparisons may be made here, with Piaget's (1952) postulated 'stages' of development. In the pre-operational stage (before seven years) the child's reasoning is dominated by perception of external stimuli (Berlyne, 1957). For example, when the pre-operational child picks coloured balls out of a short, thick, glass tube, and places them in a long, thin one, he believes that their number has increased--since the level is now higher. At the age of seven or thereabouts, the child acquires the ability to perform 'concrete' operations, i.e., 'reversible internalizable actions'. These actions--serial ordering operations, formation of classes, establishing relations between classes--can at first be performed only with objects in the immediately present environment. But by the attainment of 12 years, a stage is reached in which the child performs formal operations, reasoning with propositions and hypotheses.

The gradual emergence of thought as independent of sensory perception is in harmony with what has been said above. In the concrete operation stage, mental organization is not sufficiently complete to provide sufficient attention-cathexis for independent operations (as in the 'pebble' example). Consequently, the additional excitation of external sensory stimuli is necessary to initiate the sequence. In the formed stage, however, the adolescent is able to apply formal operations to what are, essentially, class representatives.

[17]For example, most people, even after watching the construction of an Euler one-sided surface--i.e., a surface, every point of which can be transversed by a single line, are unable to visualize the motion of an object around it until they have actually traced out such a locus with their fingers.

Types of Mental Operation

For our purposes, two principal types or modes of mental operation will be distinguished.

Type A

Operations performed on representatives will be designated 'Type A'. They may be single operations (as in the pebble example), or they may be compounded, in which case, some kind of 'storage' must intervene. Consider, for example, the following "directions" questions from the Stanford-Binet scale: 'A man walks one mile west, then one mile north, then one mile east. How far is he from his starting point?' The question is attempted without the aid of pencil and paper.

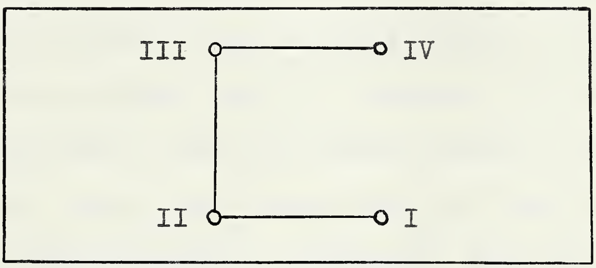


Fig. 3. Successive operations in storage.

An individual attempting this question mentally, begins by activating the series of phase sequences

corresponding to movement I⟶II.[18] After a slight delay, (to read further instructions), the individual takes the result, which has meanwhile been 'stored' in position II, and initiates phase sequences corresponding to movement II⟶III. If the delay at position II is sufficiently lengthy, the stored reference point (the result of the first operation) may be lost, in which case, the initial phase sequence I⟶II must be reactivated. However, the reactivation time of I⟶II would be less than when the operation was first performed, so that the individual would progress beyond point II (to III, let us say). Thus, by a rapid sequence of operations, storage, and recycling, the individual is able to visualize a sequence of operations each of which is performed on the result of a previous one.

It is in order to ask at this point, if the number of steps in the problem could be increase d to the point where the individual could no longer achieve a solution. Theoretically, there would seem to be no limit as long as a supply of hypercathexis is available, for, in each recycling, the circuit resistances are gradually removed until a stage is reached where the results of the first operation

[18]The phase sequences, essentially electrical circuits, would have a potential which decays with time until the potential of the 'image' of the transposed element falls below some critical limit. The interval between the initial activation and the falling below the critical limit, can be considered as 'storage time'.

are permanently stored. Presumably at this point, the individual could begin his recycling operation at II, then at III, thus gradually attaining a starting point further along the series. (This is actually what happens when the results of an operation are given a special name or symbol). However, artificial limits tend to be 'induced' for reasons which will be explained in a subsequent section.

Type B

When the organism moves away from direct interaction with the environment, it creates for itself a region of activity in which it may exercise choice. The phenomenon of choice presents one of the most intriguing and yet elusive ideas in contemporary psychology--an idea which has resulted in scores of models, and of course, the inevitable controversies. It will be profitable to distinguish two general types of choice situations: 1. utility, and 2. non-utility.

1. Utility Decisions. The concept of utility is difficult to define, and while some writers (Edwards, 1951) accept it as an undefined term, it is informative to view utility from the point of view of the homeostasis model.

The utility of an object refers to its ability to bring the overall displacement of need variables closer to equilibrium--the greater the movement toward equilibrium, the greater the utility. However, since more than one

variable may be influenced by the given object, the problem arises as to how one should combine the displacements of two or more variables. This gives rise, in turn, to the problem of a scale for rating utilities. Specifically, is it possible to quantify the extent of the preference of an individual for object A over object B (cardinal utility) or can we merely say that A is preferred to B (ordinal utility)?

It would appear that at a specific time, an individual possesses a hierarchy of needs. When he is to choose between two objects (let us suppose it is a choice between two foods) he 'imagines' the consequences of the possession of each. This imagining amounts to a low-potential activation of the phase sequences corresponding to a real act of eating, and is followed by a similar activation of the phase sequences corresponding to the satisfaction derived from eating, i.e., the toward-equilibrium displacement of need variables. Thus, by imagining the choices in turn, the individual is able to perform what might be called a 'miniature' displacement of existing need variables. We can visualize at a choice point a rapid summation of the displacements following each imagined choice.

The psychological difficulty of a choice lies in the fact that the results of the summation resulting from each possible choice must be simultaneously compared. Since

a summation exists as an activated phase sequence whose potential decays with the passage of time, the simultaneous comparison can only be approximated to by a rapid alternation between each choice.[19]

2. Logical Choice. In a logical choice situation, the individual again faces a series of alternatives; however, the alternatives themselves do not, in this case, have a utility value, but may be regarded as means to an end which does possess utility. An example would be a rat trying to choose between paths in a maze which leads to food. When a logical choice is encountered in a mental problem, the problem solver performs mental operations in which he weighs the relative likelihood of a particular choice leading to the utility object. As in the previous case of utility choice, likelihoods must be retained and compared, a problem which presents difficulties when they are nearly equal.

---

[19]This would be true if the alternatives were of approximately the same attractiveness; choice becomes easier as the difference between the 'summations' representing each alternative increases. In this case, each alternative could be compared at times separated by a finite time-interval.

So far, only situations in which the individual is certain to obtain a known incentive ('riskless' decisions) have been considered. In the more general type of 'risky' decision, the separate alternatives have different possible outcomes, each with their own utility and expected frequency. In this case, the assumption of 'rational' behaviour leads to the hypothesis that the individual will choose the laternative with the high expected utility. (Marshack, 1956, Edwards, 1951).

This type of decision is of major importance in the discussion of mathematical ability and its elaboration is left to a later chapter.

In both utility and logical choices, there is competition between alternatives, and consequently, a division of the available hypercathexis. Not only does this distinguish Type B from Type A processes, but it also implies that the division of the available hypercathexis could, if alternatives were sufficiently numerous, cause a lowering of the amount of hypercathexis given to each alternative to the point where 'rational' choice behaviour becomes difficult or impossible. Under these circumstances, it is conceivable that the individual's choice would be made by dealing inadequately with, or by ignoring, some of the alternatives, i.e., by manifesting behaviour that can be construed as 'illogical' or 'irrational'. The proof that this does in fact, happen, is one of the major objectives of the experimental part of this thesis and lends support to the theory so far advanced.

Interaction in Thinking

In what has been said so far, we have considered the individual 'in isolation'; no attention has been given to the possible influence on thinking due to the presence of another person or group of persons. Since much of our thinking is done in the company of, and often in competition with,

others, it will be necessary to give some attention to this problem.

When organisms group together to form a society, they presumably do so because the individual is better able to satisfy his needs through interaction with the group than in isolation (Davis, 1948, p. 25). However, since the unhindered pursuit of incentives by any individual could dangerously disturb the need variables of another individual, there arises the necessity of group constraint on the freedom allowed in the satisfaction of needs. In essence, the individual attains a sensitivity to the opinion of others by acquiring the ability to see himself in the dual role of subject and object.

The child finds it to his advantage to anticipate the effect his actions will produce on others, and he gradually acquires the capacity to look at himself through the eyes of the collective potential 'others' which Mead (1934, p. 140) calls the "generalized other". This ability is internalized for the same reason that physical operations are, viz., internalization allows their performance in 'miniature' and the lack of necessity to perform a potentially dangerous act overtly. However, if the individual were completely and immediately subject to the capricious evaluations of a segment of the 'generalized other', no independent action would be possible at all. In other words,

to develop a basis for action, the individual has to have a somewhat stable concept of himself (self concept) which does not fluctuate radically in the face of what he perceives to be variations in his evaluation by the generalized other. The necessity of protecting the self concept becomes a "need" which may play a prominent role in determining limits in the mental operations considered.

<u>Energy Distributions in Reality Directed Thought</u>

The foregoing discussions may be summarized by means of Fig. 4.

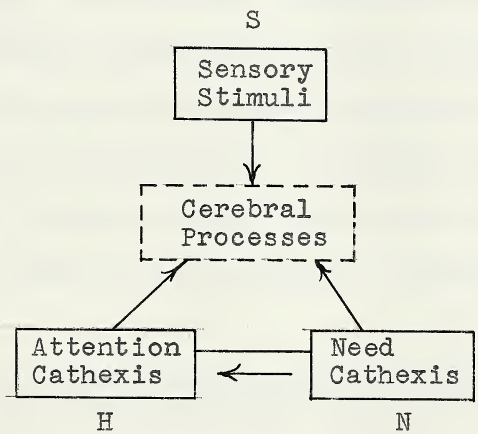


Fig. 4. Energy distributions in reality directed thought.

The energy entering into that activity of the cerebral cortex which is referred to as 'thinking' stems from two principle sources: (a) external sensory stimulation,

(b) central facilitation. The latter represents a balance between the need-cathexis and hypercathexis. Thinking, ordinarily involves some combination of the three inputs, but two tendencies may be distinguished as the human organism matures: (a) a movement away from an S-N[20] direction of thought, and (b) a gradual independence of the H source from the S and N sources.[21]

In the type of thinking processes that we have denoted Types A and B, the H source is primary but may be influenced advantageously or disadvantageously by variations in the S and N inputs. They are discussed in turn.

1. Type A. The general disruptive nature of irrelevant external stimulation need not be discussed further, but the question of needs requires elaboration. We will imagine that an individual is engaged in a Type A process. It was stated earlier that there would seem to be no theoretical limit to the number of successive levels to

---

[20]Russell (1952) constructs a gradient of thought along a dimension of diminishing influence of exterior sensory stimuli.

[21]The complete independence of centrally controlled thought does not seem to be possible. Studies in sensory deprivation (Bexton, Heron, Scott, 1954) reveal that a long term lowering of sensory inputs beyond a critical limit results in the disruption of organized behaviour patterns.

which the individual could progress. Let us suppose further that the individual's 'level-time'[22] function is given by C_1 and that he is working in a class situation where the average time for the group is given by C_2.

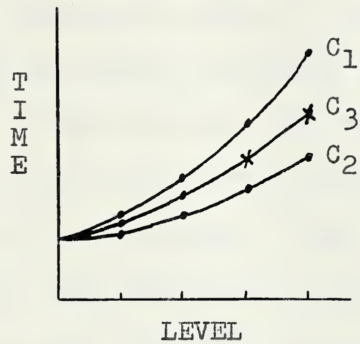


Fig. 5. Level-time curves illustrating the adverse effect of group pressure.

It seems reasonable to suppose that the individual's expectations with respect to his own performance depend to some extent on the performance of the group. If the individual is aware of his relatively poor performance,

---

[22]The curve C_1 represents a graph of the relationship between the time required to perform step-like functions (such as Type I) and the 'level' reached. Since the process consists of successive repetition of the sequence: Operation $\longrightarrow$ storage $\longrightarrow$ operation $\longrightarrow$ storage $\longrightarrow$ we may speak of each (operation $\longrightarrow$ storage) couplet as defining a 'level'.

there would arise a need to 'save face' (protect the self concept) by arriving at answers in times which correspond more closely to the group time than to the time that the individual would require to obtain a correct solution.

The activation of this irrelevant need would give rise to a partial disorganization of central control,[23] a disorganization which would presumably increase as the individual falls further behind. One way in which he could resolve this difficulty would be to exhibit behaviour which might be interpreted as guessing. This would result in a curve C_3, where the last two points represent incorrect answers. Thus the upper bound[24] evidenced in the individual's performance would be caused by the situation and should need not be construed as the lack of ability to perform the task in question.

3. Type B. Here again, the H-source is of paramount importance, but the influence of the 'needs' variables may be of positive or negative value depending on whether we are dealing with a utility or a logical situation. In a utility

---

[23]The loss of central control would follow the decrease in hypercathexis which results from the activation of an irrelevant need (i.e., irrelevant to the task to which hypercathexis is directed).

[24]An upper bound refers to a point or level above which the individual's performance is inadequate.

situation the total displacement of need variables defines utility so that an increase in a particular need may actually make a decision easier. In the logical choice situation, the incentive does not reside in the choice itself and an increase in intensity of need would deplete the available hypercathexis thereby lowering the limits which already seem to exist.[25]

---

[25]In this case, we would expect to find that a high need (i.e., strong drive) group, having less available hypercathexis, would show greater tendency to give insufficient attention to the available choices--whereas a moderate-need group would display greater wavering or uncertainty. These are the results reported by Bruner et.al. (1957).

CHAPTER III

THE PSYCHOLOGICAL BASIS OF AXIOMATIC MATHEMATICS

In considering the problem of a definition of mathematics we have to distinguish between the philosophical question concerning its proper domain (Hamley, 1934) and the psychological question concerning the mental functions or operations required in its performance. This thesis deals explicitly with the latter problem--although the two questions are not unrelated.

A survey of the history of mathematics (Smith, 1929) reveals a startling parallel between the development of mathematics as a system of knowledge and the maturational development of the brain. Considering algebra and geometry separately, each at first passed through a stage of concrete operations (in Piaget's language), where counting was performed on the fingers and measuring involved comparison of lengths with the hand or foot. At this stage, mathematical thinking could be carried on only in connection with an immediate environment, and was therefore completely under the domination of sensory stimulation. Through a long, slow process, algebra and geometry gradually gained independence

from perception of the external world,[1] becoming internalized as a set of operations performed by the brain on 'abstract' elements. In short, that 'axiomatic' method which characterizes modern mathematics (Robinson, 1946, p. 5) represents a transition to the formal stage in which operations are performed on 'representative elements' (e.g. point, line) through a central process powered by hypercathexis.[2]

---

[1]Rignano (1923, p. 146), states:

"Thus arithmetical calculation is comprised solely of a series of operations, all of which have been materially performed in the past, but now, on the contrary, are performed in thought alone. The long and very slow evolution of this transition, the difficulties met with by primitive peoples in reaching the stage of counting a number of objects greater than the number of fingers on one hand, or two hands, or of fingers and toes together, the slow evolution of the abacus and its persistence even into relatively recent times, the late, and for so long, excessively imperfect compilation of the mnemmonic 'tables' of addition, and later of multiplication, and the very slow evolution of the calculus of fractions, all this shows how very great were the difficulties which man had to overcome in empirically discovering one by one, the results of certain operations of calculation, before he was able to reach the stage of performing the operations in thought alone."

[2]It is an interesting historical fact that the axiomatization of geometry preceded that of algebra by some 2,000 years. However, in some respects, Euclid's work is still representative of the 'informal' stage, for these imperfections which were later to be challenged (Veblen, 1904),--congruence by superposition and the celebrated 'fifth' postulate are cases in point--are clearly due to the domination of perception of the external world. The complete axiomatization of geometry and algebra took place about the same time, viz., in the latter part of the 19th century.

The formalization of mathematics undoubtedly ranks as the greatest single event in its history. It is, however, to be understood not as a historical accident, but as the inevitable consequence of the mind's attempt to free itself from the domination of its environment. The thesis that the mental functions utilized in axiomatic mathematics consist of 'operations' performed on <u>representatives</u> will be examined further in the following section.

The Structure of the Axiomatic System

In order to have a point of reference for the following discussion it will be useful to examine a specific set of axioms (Appendix A). Although the axioms for projective geometry were chosen (Robinson, 1946) we could equally well have used those of symbolic logic (Hilbert and Ackerman, 1950), general topology (Sierpinski, 1952), Euclidean geometry (Veblen, 1904), non-Euclidean geometry (Woods, 1904) or of any other branch of modern mathematics.

The following components of a modern mathematical system are illustrated in these examples:

1. <u>Undefined elements</u>. Inherent in the chain of deductive argument is the necessity of starting somewhere; thus an axiomatic system begins with classes of undefined elements. It has previously been pointed out that the psychological necessity of dealing with classes results from the many-to-one mapping of the environment on the brain.

In geometry, the undefined elements are point and line. These geometric entities have the special property that they retain their identity independently of one's point of viewing them. Thus they have enjoyed a special role in the mapping of the environment on the brain. In particular, the central processes are able to activate low-resistance representatives of these classes.

In algebra, we began with a class of elements A, B, C, without specifying their identity.[3] However, from both a historical and mathematical point of view, these classes of elements can be identified with such entities as numbers, points, or lines, with which the individual has had previous experience.[4]

---

[3]Mathematicians commonly speak of 'abstract' elements or the 'abstract' nature of mathematics (Rosenbaum, 1958). This notion needs some qualification. A symbol x is abstract in the sense that it can be a representative of any one of many different classes, or of one of many elements within one class, but there is a suggestion in the implication that "mathematics is independent...in some sense... of empirical knowledge" (Kenner, 1958) which seems contrary to the principles of operation of the brain. An abstract symbol must be attachable to some class--i.e., to some neural locus--or it could not be mentally represented and comprehended at all.

[4]After dealing with the concept of class over a period of time, the individual may adopt a representative class. Thus upon reading the expression: "let a, b, c, ... be a class of elements..." many people immediately picture a vague cluster of dots.

2. Defined terms. Definitions are statements of relations which hold between the term to be defined and the 'represented' undefined terms. In projective geometry, for example, a 'plane' is defined in terms of lines, points, and the relation of 'lying on' and 'passing through'.

2. Axioms. The axioms are accepted as true statements (Allandoerfer, 1957, p. 69) about the defined and undefined elements[5] from which further true statements are to be generated. The choice of axioms is completely at the disposal of the system constructor except that an attempt is made to ensure that the entire system of axioms possesses the following properties: 1. independence, 2. consistency, and 3. completeness (Hilbert and Ackerman, 1928, p. 34).[6]

---

[5]Some mathematicians distinguish between axioms as "statements of fact" and postulates as "conditions to be satisfied" (Huntington, 1904, p. 172). This distinction, which has existed primarily in geometry, is apparently disappearing, leaving the word 'axiom' to refer to any set of conditions adopted as the basis of an abstract science.

[6]The criterion of 'independence' of the set of axioms reflects the desire to reduce the set to the smallest possible number of axioms. 'Consistency' is imposed so the system will not generate self-contradictory results. 'Completeness' is a more difficult concept, but it means essentially that any statement which is phrased in the 'language' of the system can be proven to be true or false by a proof which is itself entirely phrased in the language of the system. The latter condition has proven to be difficult to achieve--as the celebrated Gödel therem illustrates (Katsoff, 1949).

The axiomatization of mathematics represents a pure

From a psychological point of view a necessary condition for the comprehensibility of the axioms is that they must be verifiable by reference to some class of objects previously experienced (and therefore capable of mental representation). Thus, while the associative law of algebra may be written:

$$a o (b o c) = (a o b) o c,$$

the individual can be said to comprehend this statement only because he is able to verify it by reference to specific representatives of a class. In this case, for example, he may substitute integers for a, b, c, and the arithmetical (+) for the undefined operation 'o'.

4. Operations. When the axioms or true statements have been specified, some means of combining true statements to generate new ones must be provided. These rules take the form of operations to be performed on the given elements. These may be formalized by definition, such as the addition

---

form of model construction. A complex system in the real world--the real number system for example--is mapped on a structure consisting of five axioms (Peano's axioms). Starting from these axioms--and without further reference to the real-world system--the mathematician through a series of mental operations can reconstruct all the properties of the number system. In other words, he has followed the basic mode of operation of the brain by mapping a complex segment of the exterior environment into the smallest possible number of elements in terms of which it can be comprehended.

(+) and multiplication (•) of number theory (Robinson, 1946, p. 74) and the "rule of substitution" of the calculus of propositions (Hilbert and Ackerman, 1928, p. 28) or they may be informal implied operations as the ordinary rules of combination of the propositional calculus are implied in projective geometry.

5. System of Theorems. The application of the permissible operations to the original statements (involving the undefined elements) generates a sequence of 'true' statements. The more important ones, usually those which will be used as reference points, may be labelled 'theorems'.

The list of essentials of a modern mathematical system is now complete. From a psychological point of view, an axiom system begins with representative elements which are subjected to certain operations. Each mental operation represents the reactivation of a phase sequence (through a central process powered by hypercathexis) which is, in fact, the motor component of a neural locus established during the performance of a physical act.[7] Mathematics represents a pure form of attention-cathexis directed thought in that its

---

[7]Rignano devotes three chapters of his book, The Psychology of Thinking (1923, p. 141-208) to a historical and psychological demonstration of the thesis that mathematical operations represent internalizations of physical actions, which could be performed overtly but are

'representatives' (point, line, element) are devoid of specific content.

---

performed mentally instead. This work is in accord with Piaget, who considers "mental operations" to be "internalized physical operations". The physical nature of such important mathematical ideas as ordering, corresponding, function, and the taking of limits is obvious (Hamley, 1923). Moreover, the concept of an 'operation' as an action performed on an element can only have psychological meaning as a type of movement, which, in turn, means that an operation can only be represented through an activation of a phase sequence corresponding to some movement which is performable externally.

CHAPTER IV

LEVEL ONE ABILITY

The student of mathematics who progresses far enough to encounter the axiomatic scheme described in the last chapter may exhibit various types of performance in relationship to this structure. Mathematical ability, then, could be defined in terms of relative performance in one of these tasks. Consequently, it is not unreasonable to speak of different <u>kinds</u> of mathematical ability--corresponding essentially to different modes of mental operation.

From the first encounter with the axiomatic method in the high school, through to graduate courses in pure mathematics, the student's major role could be described as <u>comprehension</u> of a system already developed. It is important at this point to investigate the psychological significance of the word <u>comprehension</u>.[1]

---

[1]The finished axiomatic system consists of a chain of theorems emanating from the axioms. Passage from one link to the next is <u>demonstrated</u> by deduction--a special form of which is the syllogism:

$$\begin{array}{l} P \longrightarrow Q \\ \underline{P \qquad\quad} \\ \therefore \quad Q \end{array}$$

However, there is no necessary correspondence between the exhibited steps of the deduction and the underlying psychological processes by which the conclusion is reached. Under the circumstances it may be wise not to refer to deduction as a mental process or a form of reasoning, but

Initially, the comprehension of the axioms means that they can be enacted on the representatives of some class known to the individual. Thus the 'commutative' law of algebra

$$a \circ b = b \circ a$$

has meaning[2] for a student because he knows that he can substitute the ordinary integers and the operation '+' to get

$$3+2 = 2+3$$

Similarly, the student of projective geometry comprehends Axiom IV because 'point' and 'lines' are elements for which he has representatives available, and the mental manipulation of points (or lines) is part of his mental

---

rather as the method of exhibiting the results of reasoning. This point of view was put forth by Binet (1907) and is reiterated by Vinacke (1952).

Induction may have greater claim to the status of a genuine psychological process. Essentially, it involves the attribution of some property of a member of a class to all members of the class. This would seem to be closely related to the process of division of the external environment into classes.

[2]According to Carnap (1942), an individual can be said to comprehend the 'meaning' of a symbol if he knows where it applies. Conversely, he would not know the 'meaning' of the symbol unless he knew at least one place where it applies.

repertoire of internalized operations.[3,4]

This kind of basic comprehension, which involves imagined operations performed on primitive elements and which gives rise to the subjectively felt 'truth' of the axiom, will be denoted primary reference.

The development of an axiom system could conceivably proceed by primary reference alone, but it usually does not do so. To understand the reason for this, we will consider the demonstration of the solution of the following simple problem to a grade nine student who, as yet, has no knowledge of algebra:

When three is subtracted from a certain number, the result is eight. Find the number.

---

[3]An 'internalized' operation is one which has been performed externally sufficiently often that it possesses a well established neural pattern capable of low-potential internal excitation.

[4]Many students report that while trying to picture a line, they experience a feeling of motion, as though the line were actually being drawn before them. Again, most people do not seem to be able to imagine more complicated geometric figures (e.g. triangles), without imagining a line being traced around the outside. The perception of a triangle is not a static affair, but involves the rapid traversing of its circumference (Hebb, 1949, p. 85) so that the feeling of movement during the reactivation of this perception may be due to the physical movement of the eyes, or the hand as in tracing it, or both. Consequently, the activated neural components of physical action may vivify static as well as dynamic images.

The demonstration of its solution readily lends itself to primary reference. For we can picture (even draw perhaps) a set of elements--and three removed. We then have two clusters, one containing eight elements, the other, three elements. Physical replacement of the three elements gives a set of eleven.

When the student has mastered some elementary algebra, a solution can be demonstrated as follows:

$$X-3 = 8$$
$$X = 8+3$$
$$X = 11$$

The demonstration is now different in this respect; that while there is still physical manipulation (of the symbols), there is no longer manipulation of the primitive elements (points) of the system. The manipulation now proceeds by means of a set of rules, devised for the sake of economy. A list of rules (as supplied to a grade nine student) would be:

(a) All the terms containing the unknown are collected on the left hand side of the equation. If a term is 'transposed' across the equal sign, then the sign is changed.
(b) All the terms not containing the unknown are brought to the right hand side.
(c) The terms on each side are added algebraically.
(d) The number on the right hand side is divided by the co-efficient of the unknown term.

We can, of course, solve this problem mentally by picturing the equation, and by imagining the physical transposition

of the terms. This mental act (which represents the internalization of the physical act of transposition), again obtains its 'correctness' by reference to a physical act which may be performed. But since this physical reference act proceeds according to an arbitrary set of rules, introduced for the sake of economy, rather than by the operations performed on the primitive set, we say that the subjectively felt 'truth' of the mental act is established by <u>secondary reference</u>.

Secondary reference--and the successive consolidation of rules or results--allows mathematical systems to develop along a vertical dimension. For once a result is established, it acts as a <u>reference point</u> and abnegates the necessity of returning to the axioms as a starting point[5] for subsequent demonstrations. Reference points need not be expressed as a set of rules. The theorems of any axiom system--as for example, the propositions of Euclidean geometry--act as reference points for the further development of the deductive chain.[5]

---

[5]This is a major reason for the power of mathematical thinking, as well as a source of its difficulty. A student may learn historical facts, or new words, by a process which is essentially horizontal--one fact not necessarily relying on the previous one. In mathematics, however, the loss of a reference point may render the deductive chain unintelligible from that point onward.

The student may encounter two situations with regard to comprehension of a system:

(a) He may be supplied with a complete demonstration of an argument. The simple example above will suffice to illustrate this point. Here the student's role is principally <u>verification</u> (by secondary reference to the rules or results of the system). If the operation which leads from step n to step n+1 of the development is stated specifically (e.g., "transpose"), then the individual performs this act mentally and verifies the result by checking it against the recorded result. If the operation between the steps is not specified, however, then the individual must locate it (from among the permissible operations and results of the system), apply it, then verify mentally.

(b) In many computational situations, the rules are sufficiently well formulated that the individual can perform them himself. This is the case in all of the high school algebra as well as in a great part of differential and integral calculus, solution of systems of equations, and so on. In this case, the individual is convinced of the correctness of his argument because he has performed a series of manipulations according to the set of rules, and he has previously established (or has seen established) the correctness of these rules (again through primary or

secondary reference).[6]

It will likely be agreed that the type of ability demonstrated in situations (a) and (b) is of a necessary, but rather elementary nature. For the purpose of having a convenient label, we will use the term Level I ability to refer to the individual's relative performance in these types of tasks. From the point of view of mental functions, Level I ability, which involves a set of operations performed on representatives, the storage of reference points, and further performance of operations on the reference points, requires the type of mental performance previously described as Type A.

Upper Bounds in Level One Performance

The dependence of Level I ability on Type A functioning suggests that the former may exhibit the induced upper limits which are found to exist in the latter. The obvious difficulty which many students experience in their attempts to comprehend mathematics warrants a closer examination of forces which determine these upper bounds.

---

[6] Of course, the bane of the rule-application method is that a student often acquires rules which are faithfully applied without being understood. Nevertheless, even this kind of performance can be useful when a purely computational end is in view.

1. Natural Upper Bounds. A prerequisite for the operation of Level I ability is a supply of hypercathexis which is sufficient to fix successive reference points. This reaches particularly large proportions when a reference point involves a 'condensed' symbol (Rignano, 1923, p. 16). This can be well illustrated by the notion of the limit of a sequence.

By definition, a sequence X_1, X_2, . . . is said to possess the limit L if subsequent to the choice of any number e, however small, one can find an integer N such that $|X_n - L| < e$ for all values of $n > N$ (Sokolnikoff, 1939, p. 6).

The dependence of this notion on a sequence of physical operations is easily demonstrated. We suppose that the series $\{X_i\}$ can be represented by points on the line in Fig. 6.

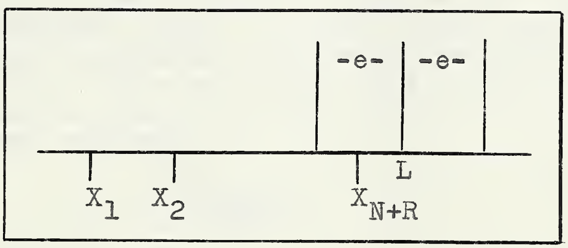


Fig. 6. Diagrammatic illustration of the limit point of a sequence.

We take a small length e, and construct the interval $L \pm e$. When we try to visualize the existence of a limit point

(according to the above definition), we imagine the physical 'drawing' of the sequences, moving from left to right, until point after point falls within the small interval. (The author visualizes a series of vertical strokes being drawn on the line.) Thus condensed within the symbol $\lim_{n \to \infty} X_n$ is an infinite number of implicit physical movements.[7]

The consolidation of such reference points requires a considerable expenditure of hypercathexis. As we proceed up the inverted-pyramid structure of mathematics, the number of reference points increases, since every new theorem adds one more to the accumulated list. The hundreds of definitions and theorems in a book such as Sierpinski's General Topology (1952), coupled with the fact that the later theorems require a storage in memory of the results of

---

[7]The condensation principle gives to the differential calculus its tremendous power, ('power' being defined by the number of physical operations which would be needed to replace the symbol). When we recall that a result dy/dx of an infinite number of operations can itself be used as an element in another sequence

$$\frac{d(dx/dy)}{dx}$$

then we begin to appreciate in some measure the economy that is introduced into human thinking through the use of the condensed symbol.

preceding theorems,[8] gives some idea of the tremendous continued supply of attention-cathexis which is necessary for a mastery of an axiomatic system. Nevertheless, the continued progress along the deductive chain which occurs under the application of a constant supply of attention-cathexis (as in Type A functioning) indicates that upper bounds in Level I performance should only appear for one of the following reasons: (a) the attention-cathexis supply is depleted through the cyclic fluctuation of basic needs,

---

[8]This does not mean that all the reference points established to the end of the n'th theorem will be used to establish the (n+1)'st theorem, but they are all potentially usable, and unless a substantial proportion of them are retained, the development becomes incomprehensible.

The perennial difficulty which students experience in comprehending mathematics lectures is based on the fact that while comprehension is probably an exponential function of deductive length (C_1), mathematicians tend to develop their arguments in a linear fashion (C_2). This inadequate allowance for storage results in a diminishing comprehension curve (C_3).

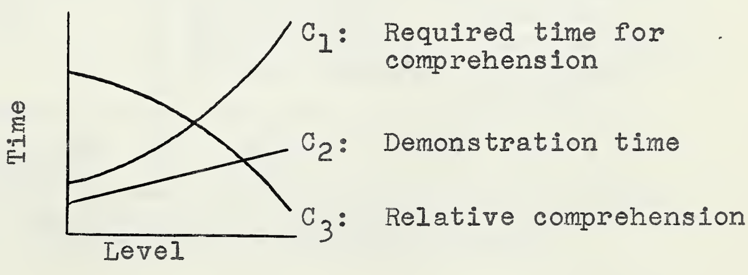


Fig. 7. The decrease in comprehension under linear demonstration time.

(b) the existence of low motivation results in a disinclination to apply available hypercathexis to the task at hand.[9]

2. Depression of the Upper Bound. In addition to difficulties inherent in the axiomatic method due to its very nature, and to the mental and motivational equipment of the student, certain other forces may act which tend to depress the upper bound of achievement in the axiomatic chain. These forces can be considered as artificial in the sense that they cannot be attributed to the student, and could presumably be eliminated under ideal circumstances.

---

[9] Motivation can be understood as a utility choice situation in which the individual weighs--against an existing need configuration--the respective utilities of a corrent solution on the one hand and some extraneous behaviour (such as daydreaming) on the other.

Choices Available

	Attempt Solution	Extraneous Activity
Utility	Satisfaction derived from correct solution	Pleasure derived from extraneous ability
Probability of Success	$p < 1.0$	$p=1.0$

Fig. 8. Motivation as a utility decision.

This is a type of 'risky' decision, for the probability of a satisfactory solution is less than certainty. As this probability decreases through repeated experiences of failure,

3. <u>The Creation of Gaps</u>. While the phenomenon of gaps is best illustrated in such highly formalized systems as the calculus of propositions where the permissible operations are specifically defined, similar argument can be advanced for any system.

Let us suppose that the system in question utilizes n operations in the development of the sequence S_N, S_{N+1}, S_{N+2}, . . . Suppose further, that each successive step S_{N+1} is generated by combining one of the S_{N-1} previous reference points with S_N, using for this purpose one of the n rules.[10]

We have seen that if the demonstration of the transition from S_N to S_{N+1} is accompanied by a statement of the operation applied and the previous result used, then the student's task is mainly one of verification. If, however, S_{N+1} follows S_N without explanation, then the student must find <u>which</u> operation was used to combine <u>which</u> statement with S_N. We shall say in this case that a 'gap' of one unit length has been created; moreover, if we assume

---

the expected utility (Siegel, 1957) to be derived from attempting the problem decreases as well, thereby increasing the likelihood that the individual will exhibit behaviour indicating a preference for extraneous behaviour.

[10]This is probably the simplest case which could occur. The algebra is deliberately approximate and useful for purposes of illustration only.

that the individual is obliged to fill the gap by combining the previous statements with S_N by a random selection among the permissible operations, then there would be $n \cdot N$ possible choices. For a gap of length L, the number of possible choices will be of the order of $n^L(N+L)!/N!$. Ignoring the crudeness of the algebra, the essential point remains that the number of possible random choices--and the subsequent difficulty for the student--increase rapidly as steps in the proof are omitted.[11]

We can now summarize discussions from this and the preceding chapters by postulating three possible reasons for the lowering of the student's potential upper bound in mathematical comprehension; (a) the effect of group pressure, (b) inadequate time to consolidate reference points, (c) the creation of gaps.

---

[11]Unfortunately, it is fashionable to present mathematical expositions in an extremely terse form, and to prefix tremendous gaps in the development with the statement "it is obvious that . . .".

CHAPTER V

LEVEL TWO - THE MATHEMATICS PROBLEM

The generation of a mathematical proof evolves a sequence of true statements S_N, S_{N+1}, . . . If a system has been developed to the stage S_N and if the student is asked to demonstrate the result S_{N+R}, he <u>may</u> complete the intervening gap by responding in a random manner. However, some situations are of sufficient importance that they tend to recur, and in this case it may be possible for the student to develop a strategy--i.e., a non-random method of procedure in the face of a sequence of events whose range is known.[1] When the latter type of behaviour occurs in conjunction with an axiom system, we shall refer to it as <u>problem solving</u> and shall assign it the label 'Level II' ability.

An example from the differential calculus will illustrate points to be discussed subsequently (Phillips, 1927).

---

[1]The definition of strategy employed here represents a partial generalization of the 'strategy' of the theory of games (Girshick, 1954). In the present usage the term 'strategy' will imply: (a) the repetition of certain types of events, (b) the awareness on the part of the subject of the possible range of events, (c) the formulation by the subject of a more or less clearly defined plan to deal with the occurrence of a particular event in a specific manner.

The problem reads as follows:

"A rectangular sheet of paper is 6 inches wide. It is folded so that the bottom right hand corner touches the left edge of the sheet. Find the longest length that the crease may have.

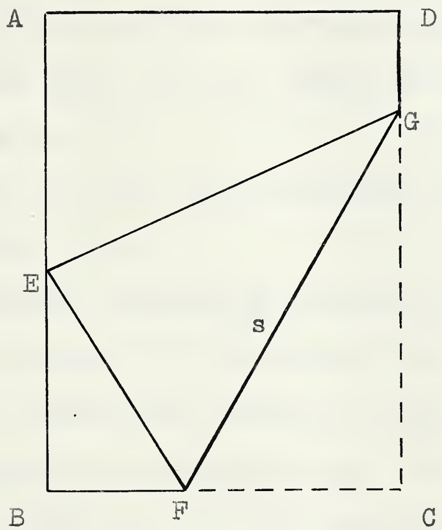


Fig. 9. Illustrative problem from the differential calculus.

We may first remark that the student must have some knowledge of the calculus[2] before a solution is possible. If, in addition, the question is recognized as a 'maxima-minima' problem, then the student would have available a

---

[2]In particular, the student must be acquainted with the theorem which provides the means of determining the maximum and minimum values of a function.

strategy which may be formulated in the following terms:

(a) Select a variable α, the variation of which is associated with the variation of s.

(b) Express s as a function $f(\alpha, k)$ of this variable and numerical constants only.

(c) Set $d\left[f(\alpha, k)\right]/d\alpha = 0$, solving the resulting equation for the value $\alpha = \alpha_0$ which gives the maximum or minimum value of s.

(d) Find $f(\alpha_0, k)$, thus obtaining the maximum (or minimum) value of s.

The important difference between the 'rules' (above) and the rules of Level I is that while the latter specify a direct set of operations leading to S_{N+k}, the former generates a <u>sequence of choice points</u>. Choice is the essence of the problem situation. In the problem above, the application of rule (a), for example, leads to a choice involving the location of α. If we take α to be an angle, there are 9 possible angles (viz., three at E, three at F, and three at G) which may vary as s varies. As the student acquires experience, in this type of problem, he does not remain indifferent among the possible alternatives. Experience may have taught him that angles adjacent to s and involving original dimensions of the diagram are more likely to yield solvable equations upon differentiation than other angles.

In this case he would favour $\angle GFC$ and $\angle FGC$.[3] If, in a given set of problems this choice of angle (for the sake of a name we will call it an 'adjacent-original' choice) does in fact lead more readily to a solution than other choices,[4] then favouring this choice will give an advantage to the problem solver over a second problem solver who makes his choices at random.

If we continue the problem and try to express s as a function of α and the given constant 6, we encounter further choices in the method of procedure. However, as this provides no new developments, we leave the problem here.

The solution to this problem moved forward from a known theorem (concerning the value of the derivative of a function) to the result. Often, however, the problem solver moves backward from the desired result S_{N+K} to some known theorem S_N.[5] The situation is represented schematically in Fig. 10.

---

[3]We could say that his 'subjective probabilities' for these alternatives were high (Cohen, Dearnaly, Hansel, 1957).

[4]The 'adjacent-original' choice would be said to have high objective probability.

[5]In many problem-solving situations the student works in both directions, balancing an analysis against deduction.

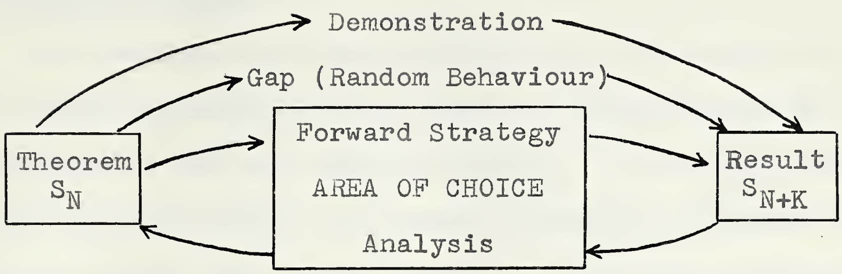


Fig. 10. Diagramatic representation of gap-filling procedures.

From a consideration of the above problem, and a wide variety of others from elementary, secondary, and advanced mathematics we can abstract a set of conditions under which problem solving may be exhibited:

(a) The individual is able to perform the set of operations which are permissible within the system and has an adequate knowledge of the reference points in the system.

(b) The individual has a strategy, a general method of attack or procedure.

(c) One or more steps of the strategy generate choice among alternatives.

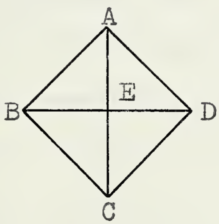
(d) The individual possesses a set of subjective probabilities which represent the extent to which he favours one alternative over another.

(e) Each alternative has an 'objective' probability--i.e., a relative frequency with which it leads to success in a set of problems of this type.

<u>The Geometry Problem</u>

In keeping with the objectives of this thesis, we will give special attention to the type of problem which is encountered by the high school student. The algebra course in the secondary school is largely devoted to computational facility through the use of rules. In Euclidean geometry, however, the student encounters the first mathematics course which deals largely with the solution of problems.

We consider the following problem which is presented to a student some time early in the grade 10 geometry course.



Given: AB = AD
BC = CD

Prove: BE = ED

Fig. 11. Geometry problem 1 illustrating uncertainty analysis.

We will also suppose that the problem occurs in the set of problems following Proposition IV (Book I), and that the student understands, in a Level I sense, the following reference points:

(T_1) <u>Definition</u>: There are 180° in a straight angle (straight line).

(T_2) <u>Proposition I</u>: If two straight lines intersect, the vertically opposite angles are equal.

(T_3) <u>Proposition II</u>: If two sides and the contained angle of one triangle are respectively equal to two sides and the contained angle of another triangle, then the triangles are congruent.

(T_4) <u>Proposition III</u>: In an isosceles triangle, the angles opposite the equal sides are equal.

(T_5) <u>Proposition IV</u>: If three sides of one triangle are respectively equal to three sides of another triangle, then the triangles are congruent.

It is possible to develop a strategy in geometry because the results D_i which the student is asked to prove constitute a finite set of the form:

(D_1). Prove: $\angle X = \angle Y$

(D_2). Prove: line AB = line CD

(D_3). Prove: $\triangle ABC \cong \triangle DEF$

(D_4). Prove: line AB $\parallel$ line CD [6]

At each choice point the individual will have to choose between the alternative methods (T_i) of establishing the type of fact under question (D_i). We will suppose that the

---

[6]Other problems can be reduced to these. For example, D_s: Prove line AB = 2(line CD) is reduced to D_2 by dividing AB in half or by doubling the length of CD.

objective frequencies for each alternative in the given set of problems can be represented by the following table. We will assume that the possible alternatives for each D_i are known to the problem solver[7] and that the subjective probabilities may or may not be equal to the objective probabilities.

Table 1

Objective Relative Frequencies in the Uncertainty Analysis of Problem 1

Type of Fact	Proposition Used				
	A	I	II	III	IV
line = line	-	-	1.0	-	-
$\triangle \equiv \triangle$	-	-	.6	-	.4
$\angle = \angle$	.2	.1	.3	.3	.1

The strategy involves an analysis which follows from the repeated application of this sequence of questions:

(Q_1) What kind of fact am I to prove (e.g. line=line)?

(Q_2) What propositions can be used?

(Q_3) Where can this proposition be applied?

(Q_4) What facts must I have before I can use it?

---

[7]In the experimental investigation reported later, this condition was ensured by providing each student with a table

If the strategy is applied to the problem given above it generates the sequence of choices shown in Fig. 12.[8]

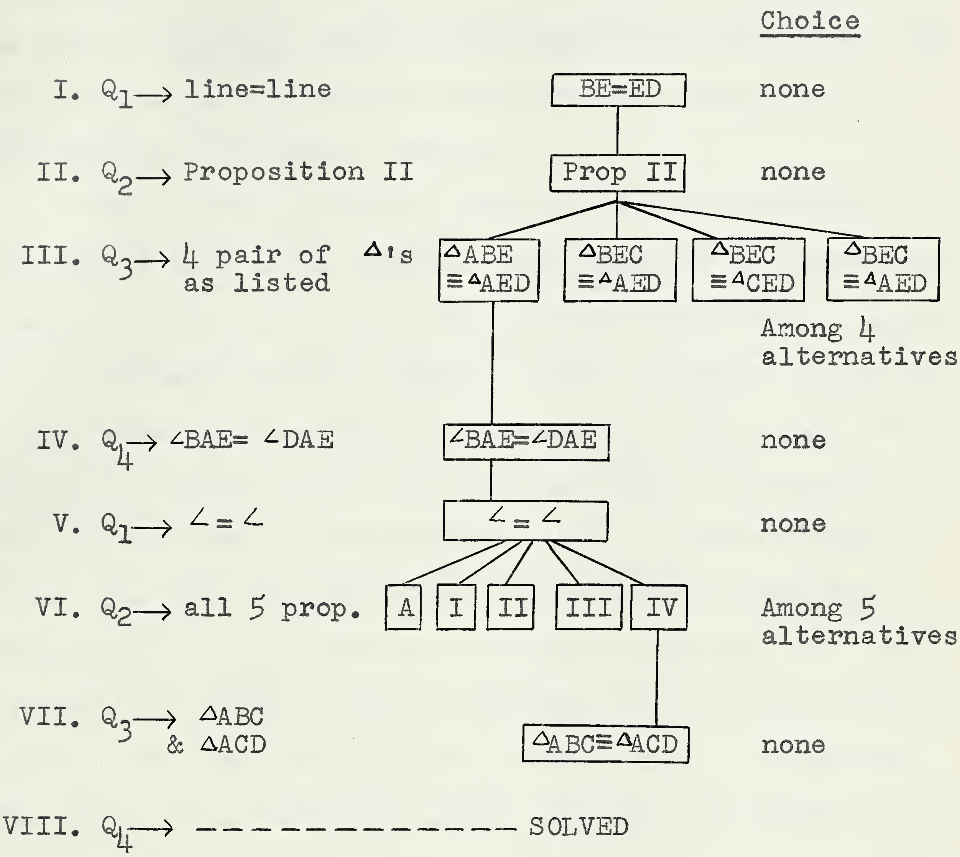


Fig. 12. Analysis diagram for geometry problem 1.

---

showing (without relative frequencies) the possible alternatives for each D_i.

[8]The symbolism Q_1 means "the question Q_1 leads to the answer ____".

The Psychological Problem of Logical Choice

The analysis diagram illustrates the sequence of choices generated by the problem-solving strategy; it shows the line of correct choices which lead to the solution. The mental operations involved in the solution seem to be considerably less direct however.[9]

The initial question leads to the identification of the type of fact to be proved. This presents no uncertainty because it is specifically stated in the 'Required'. Q_2 can also be answered without making a choice, because there is, at this point in the student's geometrical life, only one proposition which proves the required fact. Q_3 presents the problem-solver with his first choice between alternatives, for the equality of the lines would follow from proving any of four pairs of triangles congruent. At this point the individual seems to consider each of the triangles in turn, and tests the necessary conditions for congruence (two sides and a contained angle) against the diagram as it stands. Since T_2 is not directly applicable to any of the triangles, the student imagines possible additional facts

---

[9] The description which follows is based on the observations made during a period of approximately 100 hours which the author spent listening to and recording the verbal solution of geometry problems. (Details are reported in the experimental section.)

which would allow him to apply T_2 in each case. Thus, if his thinking were verbalized, it would take the form:

"I could use T_2 in $\triangle ABE$ and $\triangle AED$ if I had $\angle BAE = \angle EAD$."
"I could use T_2 in $\triangle BEC$ and $\triangle DEC$ if I had $\angle BCE = \angle ECD$."
"I could use T_2 in $\triangle ABE$ and $\triangle CED$ if I had $AE = EC$, $AB = CD$, and $\angle BAE = \angle ECD$."
"I could use T_2 in $\triangle BEC$ and $\triangle AED$ if I had $AE = EC$, $BC = AD$, and $\angle BCE = \angle AED$."

Thus the various alternatives are not dismissed at once, but each retains a possibility of application subsequent to the proof of some additional fact or facts.[10] Since none of the additional necessary facts are directly available, the chain of verbalization must be carried further. Thus the argument proceeds: "I could use T_2 in ___ if I had ___ which would require ___." Thus there is competition between alternatives and a resulting division of the attention-cathexis.

It is interesting to offer a conjecture as to what will happen when the alternatives become so numerous, and the chain so lengthy that the available attention-cathexis is depleted below an effective level of operation. It would seem that the individual in these circumstances must either ignore some of the alternatives or terminate the chain prematurely. In the latter case, the statement "I could use

---

[10] In the choice between triangles, the alternatives are not listed in the table; similar arguments hold for the latter type of decision.

T_2 in $\triangle ABE$ and $\triangle AED$ if I had $\angle BAE = \angle EAD$", might become: "I can use T_2 in $\triangle ABE$ and $\triangle AED$ because I have $\angle BAE = \angle AED$". Thus the chain would be terminated by assuming a fact which still remains to be proved.[11]

In those cases where the laternatives are covered by the table, the individual's behaviour is further influenced by the subjective probabilities which he attaches to each alternative. The alternatives of higher subjective probability are likely to be elaborated further and returned to more often than alternatives of lower subjective probability.

Attempts to Quantify Problem Difficulty

1. Standard Methods. The possibility of obtaining an objective measure[12] of problem difficulty is an attractive one for several reasons. At the present time, the standard procedure for ascertaining the difficulty of a problem is to administer it to a group of students and to determine a difficulty coefficient based on the percentage

---

[11]Often the individual makes the assumption with the full knowledge that it is an assumption. This kind of behaviour can sometimes be interpreted as a face-saving bluff.

[12]An objective measure would be one which can be calculated once the possible alternatives and the probabilities are known. In other words, it would depend on the structure of the problem itself, without having reference to a particular group of students.

who succeed. Unfortunately, this method is difficult to apply when the problem-solving time may be measured in hours. Moreover, the difficulty index applies only to the group in which it was obtained, so that this measure has a comparative but not a predictive value.[13] Again the group method glosses over the psychological difficulties of the problem itself.

Under the circumstances, it is not hard to understand why there is no standard for the high school teacher to follow in determining problem difficulty, except to use the demonstration (or deductive) length of the solution. Thus, the solution to the example treated above could be written:

$$\begin{array}{ll}
\text{Prop. IV} \longrightarrow & \triangle ABC = \triangle ACD \\
\text{Defn. of congruence} \longrightarrow & \angle BAC = \angle DAC \\
\text{Prop. II} \longrightarrow & \triangle ABE = \triangle AED \\
\text{Defn. of congruence} \longrightarrow & BE = ED
\end{array}$$

This method in itself does not necessarily correspond closely to the experienced difficulty of the problem because it does not follow the natural psychological development of the solution.

2. Uncertainty Measure of Problem Difficulty. A measure which follows the psychological processes involved in the solution must take account of the amount of choice

---

[13] Thus we cannot predict, for example, that student X will be able to solve three out of 5 problems at difficulty level Y, although this type of prediction is most useful when we consider segregating students on the basis of ability.

available at each decision point, or the amount of mental experimentation necessary to choose a correct alternative. One method of estimating the necessary experimentation would be to first calculate the possibility of a correct choice by an 'a priori' selection among the alternatives, using the relative frequencies of the table, but without performing mental experimentation. If the probability of a correct 'a priori' choice is high, then the amount of necessary experimentation would be low, and vice versa.[14]

For this purpose we present Table 2 in a generalized form.

Table 2

Probability Table For Uncertainty Analysis of General Geometry Problem

Type of Fact	Proposition Used			
	T_1	T_2	. . .	T_r
D_1	$P_{1,1}$	$P_{1,2}$		$P_{1,r}$
D_2	$P_{2,1}$	$P_{2,2}$		$P_{2,r}$
⋮				
D_k	$P_{k,1}$			$P_{k,r}$

[14]This method is used by Cronbach (1952, p. 9) in his development of an information measure for psychological tests.

The D_i are the 'types' of decisions that are made in geometry. For example, D_i might represent an 'angle decision', i.e., a choice among all the propositions T_1, T_2, . . . T_r, all of which can be used to prove that one angle is equal to another. The number p_{ij} represents the relative frequency with which the laternative T_j will be the 'correct' one to use in a D_i decision which is made within the set of problems to which this table applies.

If N_i represents the total number of D_i decisions in the set of problems and if n_{ij} is the number of times T_j is used to make a D_i decision, then,

$$p_{ij}=n_{ij}/N_i, \ \sum_{j=1}^{r} n_{ij}=N_i$$

Suppose a choice point involves a D_i decision which calls for proposition T_j. Prior to the performance of the 'thought' experiment, the probability that the individual could make the correct choice from a <u>knowledge of the table alone</u> is n_{ij}/N_i, and the probability that he could (by knowledge of the table alone) correctly make the n_{ij} decisions of this type which occur in the set is

$$\left(\frac{n_{ij}}{N_i}\right)^{n_{ij}} \text{ or } \left(\frac{n_{ij}}{N_i}\right)^{p_{ij}N_i}$$

$$= p_{ij}^{\ p_{ij}N_i}$$

Thus the probability of his making all the decisions of the type D_i correctly is

$$\prod_{j=1}^{r} p_{ij}^{\ p_{ij}N_i}$$

The geometric mean over the N_i decisions of the type D_i will be

$$\prod_{j=1}^{r} p_{ij}^{\ p_{ij}}$$

This may be called the 'a priori' probability of a correct D_i decision (using knowledge of table values).[15] However, before a correct choice can be made, the probability of its being correct in <u>that given situation</u> must equal unity, and the logarithm of this probability must be zero. The logarithm of the 'a priori' probability is

$$\log \prod_{j=1}^{r} p_{ij}^{\ p_{ij}} = \sum_{j=1}^{r} p_{ij} \log p_{ij}$$

and this quantity must be raised to zero at each choice point.

---

[15] This is the estimated probability of a correct single choice (Cronbach, 1952, p. 9).

The difficulty of the choice, or the extent of the mental experimenting necessary to raise the probability to zero, is, therefore, on the average,

$$0 - \sum_{j=1}^{r} p_{ij} \log p_{ij}$$

$$= - \sum_{j=1}^{r} p_{ij} \log p_{ij} \quad ^{16}$$

---

[16]Shannon (1949, pp. 19-22) develops a similar formula. He begins by asking,

"Suppose we have a set of possible events whose probabilities of occurrence are $p_1, p_2, \ldots p_n$. The probabilities are known, but that is all we know concerning which event will occur. Can we find a measure of how much choice is involved in the selection of the event or how uncertain we are of the result?"

He suggests that the uncertainty function $H (p_1, p_2, \ldots p_n)$ should have the following properties:

(a) H should be continuous in the p_i.

(b) If all the p_i are equal, and $p_i=1/n$, then H should be a monotone increasing function of the n. With equally likely events, there is more choice or uncertainty, when there are more possible events.

(c) If a choice be broken down into two successive choices, the original H should be the weighted sum of the individual values of H.

Shannon then proved that the only H satisfying the above assumptions is of the form

$$H = -K \sum_{i=i}^{n} p_i \log p_i, \text{ K a position constant}$$

This is the identical with Shannon's uncertainty measure. If a problem contains s decision points, and if we assume that the sequence of decisions is independent, then the problem difficulty can be represented by the expression

$$U_0 = -\sum_{i=1}^{s} \sum_{j=1}^{r} p_{ij} \log p_{ij}$$

This derivation assumes that the individual has knowledge of the quantities p_{ij}, i.e., that his subjective probabilities coincide with the objective probabilities. Suppose, however, that the subjective probabilities are w_{ij}, numbers which are not necessarily equal to p_{ij}. By a derivation similar to that above, we may arrive at a problem difficulty

$$U_t = -\sum_{i=1}^{s} \sum_{j=1}^{r} p_{ij} \log w_{ij}.$$

Moreover, it can be proved that $U_0 < U_t$.

We have so far dealt with the class of decisions D_i for which 'a priori' probabilities p_{ij} are tabulated. There is, however, another class of decisions that may arise where the probabilities have not been previously tabulated. In these cases, we will assume that any choice (without mental experiment) would be made strictly on the basis of chance and we will assign equal 'a priori' relative frequencies to them.

(As an example, we might consider the choice of triangles in the problem dealt with.)

The formula

$$U = -\sum_{x=1}^{N} p_x \log p_x$$

has many interesting psychological properties which strengthen its claim to be a measure of choice or uncertainty.

(a) The value of U is a maximum when the probabilities of each alternative are equal. This is in agreement with common sense--the most difficult decision is between alternatives of equal likelihood. Moreover, if there are no alternatives, and therefore no choice, U then takes the value 0.

(b) The maximum value of U increases as the number of alternatives increases. Clearly, increasing the number of alternatives should increase the available choice.

(c) The amount of uncertainty is increased if the individual does not know the objective probabilities--i.e., if there exist discrepancies between 'subjective' and 'objective' probabilities.

3. Numerical Example. We next apply the uncertainty to the example considered above, in the following cases:

Case (a). $w_{ij} = p_{ij}$

Choice at III, $C_{III} = -\sum_{m=1}^{4} p_m \log_2 p_m$

$$= -(.25\log_2 .25 + .25\log_2 .25 + .25\log_2 .25 + .25\log_2 .25)$$

$$= \log_2 4$$

$$= 2.00 \text{ units}$$[17]

---

[17]The choice of logarithms employing the base two is a matter of convenience, and stems from the extensive use which is made of binary systems in communication systems.

Choice at VI, $C_{VI}=-\sum_{j=1}^{r} p_{ij}\log_2 p_{ij}$

$=-(.2\log 2+.1\log .1+.3\log .3+.1\log .1)$

$=-(.4644+.3322+.5211+.5211+.3322)$

$= 2.17$ units

Total uncertainty 4.17 units.

<u>Case (b)</u>. $w_{ij}\ p_{ij}$

Suppose that $w_{12}=w_{14}=.5$

and that $w_{31}=w_{32}=w_{33}=w_{34}=w_{35}=.2$

Then, Q_2 will lead to a decision with

$$-\sum_{j=1}^{2} p_{ij}\log w_{ij}$$

$C_{II}=-1.0\log .5 - 0\log .5$

$= 1$ unit of choice.

$$C_{III}=-\sum_{m=1}^{4} p_m\log_2 p_m$$

$= 2.00$ units as before.

$$C_{VI}=-\sum_{j=1}^{5} p_{ij}\log w_{ij}$$

$=-(.2\log .2+.1\log .2+.3\log .2+.3\log .2r.1\log .2)$

$=-(1.0\log_2 .2)$

$= \log_2 10 - \log_2 2$

$= 2.32$ units.

$\therefore$ Total amount of choice 5.32 units

This case shows clearly how an ignorance of the probabilities can increase the difficulty of the problem. Thus, a 'line decision' increases in difficulty from 0 units to 1.0 unit and an 'angle decision' increases in difficulty from 2.17 units to 2.32 units. In this, the model shows agreement with our psychological intuition.

The utilization of the choice or uncertainty formula is intended to give a measure of problem difficulty which corresponds more closely to the psychological processes entering into the solution than does the deduction measure. On the surface the formula seems to have a number of desirable properties which are in agreement with our psychological intuition. The psychological validity of the formula, however, will be tested in comparing the <u>theoretical</u> difficulty of a problem with the <u>actual</u> difficulty experienced by the student in attempting to solve it.

CHAPTER VI

CREATIVE ABILITY--LEVEL III

A third function which may be performed within the axiom system--and one which the mathematician will undoubtedly regard as true mathematical ability--is the <u>creative</u> function. This may be demonstrated in many ways--of which the following are illustrative: the postulation of a set of axioms for an existing system (Peano), the correction of faults in an existing system (Veblen), the modification of axioms in a system (Gaus), the development of a new type of mathematical operation (Newton). For the most part, the creative work of the mathematician takes the form of an extension of existing systems.

Creative mathematical ability represents in a way, the ultimate manifestation of a desire--and a necessity--to comprehend an infinite environment through the medium of a finite brain. The mathematician constructs classifications, or categories, which not only encompass existing systems but which provide a basis for the interpretation of as yet unknown systems--even to the extent of leading to their discovery. Thus, in building algebraic structures the mathematician not only exhibits the logical properties of such existing systems, but provides a basis for <u>discovery</u> and <u>interpretation</u> of new systems.

The psychological basis of creative ability is, for the most part, unknown. At the present time, we do not seem to have suitable concepts to deal with creativity; moreover, it is expressed over periods of time of such length that practical experimentation has been virtually impossible.

It is difficult to find a discussion of creative ability which does not invoke Wallas' (1926) four steps. According to this theory, the solution must be preceded by a period of conscious _preparation_. The problem is then relegated to the subconscious during a period of _incubation_ in which the conscious mind either rests or deals with some other problem. Then, often with startling suddenness, a suggested solution comes in a moment of _illumination_. The tentative solution is then _verified_.

While there is general agreement with Hilgard's (1957, p. 246) suggestion that "Wallas' four steps in thinking are suggestive, but not a full scientific theory", we nevertheless find this scheme used in duscussions of creative ability by Vinacke (1952), Montmasson (1931), Patrick (1952), and others. In particular, Hadamard (1949) describes many of the creative acts of the great mathematicians in these terms.

The striking thing about creative thinking is that it does not proceed wholly under the influence of conscious

effort; there is even some suggestion that the moment of illumination comes, as it were, after attention has been 'turned off' the problem. Attention-cathexis brought to bear on a problem continually activates competing alternatives. Creativity on the other hand seems to involve the embedding of existing classes in more inclusive ones--of which the existing ones are special instances.

It is possible that the neural structure of the brain provides the basis for the physical integration of the neural circuits corresponding to concept classes. Attention plays merely an introductory role; it establishes the critical requirements and activates the necessary neural components.

Summary--Chapters II-VI

This section purports to present a summary of the theoretical framework outlined in the preceding chapters. The theory may be described as an attempt to provide a psychological basis for the study of mathematical ability, by employing a theory of mental operations which steers an independent course through the positions of Hebb, Rapaport, Piaget, Rignano, Ashby, and Bruner.

The brain, evolving and maturing as an instrument which facilitates homeostatic equilibrium, is organized under the influences of environmental stimulation and

central facilitation. The external organization proceeds through a mapping of the environment on the association areas of the brain, thus effecting a division of the complex environment into functionally equivalent classes. As a result of this organization, the organism need expend less of its equilibrium-producing cathectic energies. Part of the surplus is used for internal regulation; part is at the disposal of the ego to power its reality-testing function. The brain enables the organism to maintain long range equilibrium through its ability to visualize the results of its possible actions--to perform, as it were, mental experiments.

In 'mental experimentation', external objects can be imagined by reactivating the phase sequences normally activated in their perception; they can be mentally manipulated through the performance of mental operations--again a reactivation of phase sequences established in the original physical action. Mental manipulations are often performed on specific representations--neural loci to which the 'concept' field narrows in continued attention. The attainment of the ability to perform mental experiments through attention-cathexis is a maturational process which proceeds through states of diminishing domination by perception of the external environment. Attention-cathexis-directed thought manifests two principal modes of operation. Type A processes in their purest form involve the manipulation,

storage, and further manipulation of representatives. Type B processes involve decision or choice situations where there is a resulting competition between mental states which requires that they be separately activated.

The three-way participation in cerebral control between external stimulation, need-drives and hypercathexis is further complicated by the reflexive character of the ego, which develops a need for stable self-perception. This means that situational variables can cause a redistribution of the need-hypercathexis balance, thereby depleting the available supply of hypercathexis. Thus artificial limits may be induced in Type A performance and the natural limits in Type B performance may be seriously reduced.

Mathematics seems to have passed through stages which resemble the maturation of the human brain. The history of mathematics clearly reveals an informal-operations stage in which mental operations were closely tied to corresponding physical acts. In the development of axiom systems, mathematics entered the stage of formal operations. Different types of ability can be distinguished in terms of the students' performance within the axiom system. A fundamental type of ability is necessary for the comprehension of mathematics. It requires the performance of mental operations on the basic elements of the system (point, line, element), the storage of results, and the performance of

further operations on the results stored. This in fact, is a pure example of Type A functioning, and it carries the suggestion that mathematics has gradually evolved until it conforms with the natural mode of mental operation.

It would seem that low motivation is the only natural force tending to set an upper bound to the individual's performance in Level I tasks, but that an artificial lowering of this upper bound may be caused by (a) presentations which allow inadequate time for storage, (b) gaps in a presentation, and (c) the necessity to save 'face' in a group situation.

Frequently the student is obliged to fill in a gap in the deductive chain. If he does not exhibit a random performance of operations, but rather exhibits a strategy which causes him to make a sequence of choices between alternatives, then we describe his behaviour as 'problem solving'. This is a clear case of logical uncertainty, and some attempts can be made to quantify problem difficulty when the subjective and objective probabilities of the alternatives are known. Level II ability, requiring as it does a partition of the available hypercathexis, appears to have natural upper bounds which may again be depressed in group situations.

Finally, creative ability in mathematics involves the formulation of axiom systems or the modification and

extension of existing ones. Little is known about the psychological processes involved except that creative thinking seems to require a partial freedom from attention in which the brain performs an integration of neural fields which correspond to concept classes.

SECTION B

STUDIES AT LEVEL ONE

CHAPTER VII

PRELIMINARY DATA

At the thesis meeting in April, 1958, it was agreed that the major emphasis of the experimental work should be concentrated on the study of the solution of geometry problems and on the hypothesis that when a student faces a series of problems of increasing complexity, he reaches a predictable level, or upper bound, beyond which he is unable to deal adequately with the problem. At the same time, it was agreed that a test or tests of Level I ability should be constructed and that behaviour in this area should be investigated.

Pilot Study

The months of May, June, and July of 1958 were spent in an attempt to pursue these two main objectives. On the one hand, it was feared that the macroscopic view to be gleaned from the group studies of problem solving proposed for the autumn would gloss over the subtle psychological processes involved unless a preliminary microscopic investigation was first conducted. It was therefore thought to be advisable that a prelininary intensive study of problem-solving behaviour should be undertaken. Four potential grade 10 students were hired for this purpose, and their verbalized solutions of a large battery of geometry

problems--which included the proposed autumn battery--were recorded and analyzed. This occasion also provided an opportunity to select and reject problems for the later battery. In all, the experimenter had close to 100 hours of this kind of experience prior to the autumn investigations; the details are reported in a later section.

At the same time, various existing tests were examined, especially the now classical set proposed by Rogers (1918), and others were constructed in an attempt to discover a Level I test with a high degree of lawfulness--i.e., it was hoped that a test could be found in which 'difficulty level' and 'time' were related in a manner which might be approximated by a mathematical equation. Subjects for the latter experiments were enlisted from the graduate students and caretaking staff present during the summer session.

The Autumn Investigations

1. Population. Several factors which entered into the choice of a population will be considered in turn:

(a) The programme had to be conducted in a school in which the grade 10 geometry course was taught as a single uninterrupted unit, commencing at the beginning of the school year. This condition was not met in Alberta where algebra and geometry are mixed considerably, but the Ontario curriculum suited the proposed programme.

(b) The proposed battery for the autumn testing included individual tests which required approximately two hours of testing time per subject. After the inevitable contest between the desire for intensive study on the one hand and statistical significance on the other--the number 100 was chosen as the size of a population which might satisfy both requirements.

(c) The long-range stability of upper bounds in problem-solving ability in geometry could best be tested by studying the performance of a group of students who had had a year's previous experience and who would be covering essentially the same material for a second time. The grade 12 course ordinarily begins with a review of the grade 10 work, including, to a large extent, the same problems.

It seemed desirable then, that the population should consist of two grade 10 classes and one grade 12 class, the latter acting essentially as a comparison group. In this case however, it was important to have some knowledge of the type of instruction that the grade 12 group had received in their introductory course.

(d) The stringent scheduling of the testing and interviewing periods required that the experimenter be given considerable freedom to rearrange and interrupt regular class periods. Moreover, the tests themselves required--as will become apparent later--unusual co-operation

and enthusiasm on the part of the students.

These considerations led the author to conduct his experiments in the district high school in which he had previously been employed as a teacher of mathematics--a modern, medium-sized school situated 75 miles north of Toronto, in an area whose economy derives mainly from dairy farming and small industry.

2. The Tests. In addition to the tests constructed specifically for the purpose of this thesis, a wider variety of standardized tests was employed to give an adequate coverage of the subjects' abilities and to allow a basis for comparison and interpretation of the new battery. The standardized battery may be listed as follows:

1. Differential Aptitude Tests--Numerical Ability.
2. Differential Aptitude Tests--Space Relations.
3. Dominion Advanced Group Test of Learning Capacity.
4. Gordon Personal Inventory.
5. Holtzman-Brown Survey of Study Habits and Attitudes.
6. Kuder Preference Record--Vocational Form.
7. Minnesota Counseling Inventory.
8. Otis Quick Scoring Mental Ability Tests--New Edition--Form Gamma.
9. Revised Minnesota Paper Form Board Test.
10. Watson-Glaser Critical Thinking Appraisal.

Brief descriptions of each test, including statistical data are listed in Appendix B.

The experimental battery included the tests listed in Table 3. Each test will be discussed in turn as it is encountered in the experimental work.

Some additional non-standardized sources of information are included in Table 4. The three questionnaires may be found in Appendix B.

Table 3

Experimental Tests

Name of Test	Method of Testing	Time Allowance
1. Finite Difference Number Series, Level I	Group (main study)	70 min.
	Individual (pilot study)	40 min.
2. 'Gap-Filling' Test	Individual	45 min.
3. Geometry Problems	Group (main study)	15-20 hrs.
	Individual (pilot study)	25 hrs.
4. Number Series, Level II Tests	Individual	20 min.
5. Ordering	Individual	30 min.

Table 4

Non-Standardized Sources of Information

Name	Content
1. Ontario School Record	School marks, health record, teachers' reports from grade one on.
2. Questionnaire A	Socio-economic data.
3. Questionnaire B	Subject preferences, level of aspiration.
4. Questionnaire C	Extended questionnaire concerning variables which might influence mathematical performance.

2. Testing Programme. The autumn investigations extended over 11 weeks--from the opening of school in September through to mid-November. The group tests were administered in the school library, where standardized testing procedures could be simulated. Individual tests were administered in the guidance office, which was turned over to the experimenter for his private use during his stay in the school.

The battery of standardized tests and questionnaires required approximately 7 hours for the students to complete. The geometry tests were administered intermittently during the regular geometry periods (taught by the experimenter)

and represent from 15 to 20 hours of the student's time. The individual tests required a further two hours of testing per subject. Thus the conclusions reached in this thesis are based on data representing approximately 25 hours of testing time per subject.

Preliminary Statistical Data

Table 5 presents some statistical data concerning the experimental population. It is apparent, and natural, that the grade 12 group, as a result of two additional years of maturity and school selection, should be markedly superior in the mean scores on all tests. Some idea of the overall intellectual status of the group may be gauged from the fact that the median score for the grade 10 group centered near the 50th percentile on the appropriate norms for that grade, while the median grade 12 score usually registered somewhere above the 60th percentile on the grade 12 norms. While recognizing the dangers in applying foreign norms, it is probably safe to consider the grade 10 group to be near average, and the grade 12 group to be somewhat above average in general ability.

Matched Groups

In several instances it was desirable to compare performances at the grade 10 and grade 12 levels. However, since the grade 12 group was, to some extent, superior in

Table 5

Means and Standard Deviations of Grade 10 and 12 Groups on Standardized Tests

	Grade 10 N = 69			Grade 12 N = 37		
Test	M_{10}	SD_{10}	Percentile Rank of Median Score	M_{12}	SD_{12}	Percentile Rank of Median Score
Otis IQ	102.6	9.2		111.7	8.90	
DAT Numerical	18.5	6.70	50	24.2	6.18	60
DAT Spatial	35.6	20.6	45	51.6	21.1	60
Chronological Age	15 yrs. 2 mo.	8.82 mo.		17 yrs. 1 mo.	8.16 mo.	
Study Habits	49.8	10.6	50	51.6	10.5	52
Dominion IQ	102.2	10.4		111.9	10.6	
Watson-Glaser	51.6	8.32	45	61	10.7	74
Minnesota Spatial	38.0	8.08	55	41.2	10.8	66

natural ability, comparisons would have had little meaning. In this connection, two groups of 30 students--one at each level (Table 6)--were matched in the following way:

(a) Person to person matching on the Otis IQ.

(b) Equal group means and standard deviations on the Dominion IQ.

(c) Group equality in some lesser variables, such as, proportion of good and poor algebra students, proportion of town and rural students, and scores on study habits.

Reference to their respective norms indicates that median scores in both groups ranged above the 60th percentile on most tests. It would be feasible then, to consider the differences in performance to be due to the effect of two years of maturation and experience.

It is of considerable interest to enquire just exactly what change occurs in this two year period with respect to mental performance in general. Both groups had reached the stage of formal mental operations, so that in one way, we would expect no new type of mental function or operation to emerge. On the other hand, the mean raw score on certain standard IQ tests rises until the early twenties. However, this score reflects the result of experience to a degree which cannot be accurately determined.

The significance of differences between correlated means was calculated for the matched grade 10 and 12 groups

Table 6

Means and Standard Deviations for Matched Groups

Test	Grade 10 N= 30		Grade 12 N = 30		
	M_{10}	SD_{10}	M_{12}	SD_{12}	Significance of Difference in Means
Chronological Age	15 yrs. 1 mo.	6.48 mo.	17 yrs. 1 mo.	6.00 mo.	
DAT Numerical	21.3	6.48	23.3	6.10	Not sig.
DAT Spatial	42.8	19.1	48.4	20.6	Not sig.
Dominion IQ	108.6	7.54	109.0	8.26	Not sig.
Minnesota Spatial	39.7	6.58	39.9	10.1	Not sig.
Otis IQ	108.7	7.04	108.9	6.68	Not sig.
Study Habits	45.4[a]		46.5[a]		
Watson-Glaser	54.3	9.16	59.6	10.7	Sig. at .01 level

[a]Percentile rank of median score.

using the method set out by Garrett (1953, p. 226). As shown in the table, the Watson-Glaser is the only test which shows a significant difference; in other words, as far as mean scores in spatial and numerical ability are concerned, the two groups could be considered as samples drawn from common or equivalent populations. This problem is reconsidered in later sections.

CHAPTER VIII

LEVEL ONE PERFORMANCE AND GROUP PRESSURE

Level I ability was previously defined as the continued reiteration of the following sequence: performance of operations, storage of results, performance of operations on results. While progress in such a sequence would seem to have no obvious natural upper bounds, the theory suggested that it would be possible to induce <u>artificial</u> upper bounds in a variety of ways. One of these restrictions arises from the fact that logical behaviour--manifested in this case as Level I performance--is very much a function of situation, and the theory suggests that an individual embedded in a certain social context may prefer to be 'illogical' rather than deviate too radically from the group.

Lawfulness and Level One Performance

If the brain performs mental operations by the activation of phase sequences through a hypercathexis powered central facilitation, it is reasonable to believe that primitive mental operations will be performed--under constant conditions of motivation and attention (i.e., within a fairly small time interval)--at a consistent rate. This would be, in fact, a special statement of Spearman's

law concerning constant output. It would be a further reasonable assumption that a Level I-type task might be constructed in which 'step' and 'time' are related in a manner sufficiently definite to allow mathematical formulation. Furthermore, in view of the fact that the decay of 'storage' might well involve exponential functions--to use an analogy from the theory of electrical circuits--it would not be surprising to find that the 'step-time' function is essentially exponential in form.

After experimenting with several standard tests, the author discovered that a special variety of the time-honoured 'number-series'--constructed by means of the calculus of finite differences (Hartree, 1949, p. 38)--seemed to satisfy the lawfulness requirement. Consider, for example, the following sequence of numbers:

1, 3, 6, 12, 23, __.

The method of obtaining the number required to fill the blank is illustrated in Fig. 13.

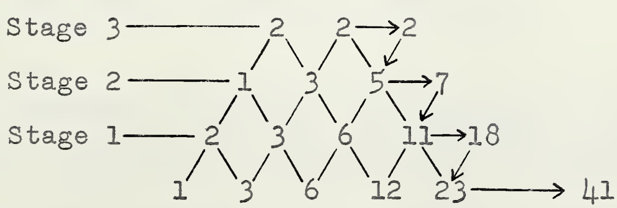


Fig. 13. A three-stage finite difference number series.

Each number in the second row (from the bottom) in Fig. 13 represents the difference between the two numbers in the first row which are immediately below and to either side of it. When situations of this kind are dealt with mentally, (i.e., without the aid of a means to record successive rows), and when the rule-of-solution is provided, the sequence of mental operations becomes: subtraction, storage, subtraction-on-results stored. This is, then, a rather simplified Level I situation. Designating each complete row as one 'stage', we could then refer to the series as a three-stage number series.

Preliminary studies of individual solutions led to the belief that this type of series possessed the useful property that the time required for solution[1] could be expressed by the equation

$$t=k_1\ e^{k2\ S}$$

Here S is an integer representing the number of 'stages' in the series, and k_1 and k_2 are constants depending on the particular individual[2] (personal constants).

---

[1]The equation represents the natural solution time for a N-stage series, i.e., the solution time under relatively constant conditions of attention and motivation.

[2]While the equation gives a good approximation to the required time for the solution of individual series of comparable arithmetical difficulty, the lawfulness of the 'stage-time' curve is greatly increased by averaging the required times of a number of such series at each level.

This equation allows us to predict the time required for the solution of an N-stage sequence, provided that the solution times corresponding to any other two series N_i, N_j are known. The equation also coincides with the theory in that it predicts no limit in Level I performance, since a series of any difficulty (say S_k) could be 'solved' provided that the quantity of time

$$t=k_1 e^{k_2 S_k}$$

was available, and that motivation toward, and attention to, the problem continued throughout this time interval.

It occurred to the author that the finite difference number series could prove to be immensely useful in the study of Level I performance in group situations, since once the constants k_1 and k_2 were determined for each subject, the 'stage-time' function would be completely determined, and deviations from the theoretical predicted values could be attributed to group influences (assuming continued motivation).

Experimental Design

1. Hypotheses and Time Limits. The experiment centered about two hypotheses:

1. H_1: The equation $t=k_1 e^{k_2 N}$ gives, for our experimental group, a workable approximation to the time t required for the solution of an N-stage series.

In order to test this hypothesis, a battery of number series was constructed beginning with one-stage problems and gradually increasing in difficulty to 7-stage problems.

2. H_2: Group pressure results in the creation of artificial upper bounds in the performance of many students. In other words, a substantial proportion of the upper bounds shown in individual performance profiles do not result from the imposed time limits or lack of ability, but rather from forces originating within the group situation.

Since the theoretical curve is determined by two points, the testing of H_1 required that each 'stage-time' curve should contain at least three points, (including one check point), i.e., that all subjects should have adequate time to solve at least one, two, and three-stage problems. Moreover, it was further desirable that some of the individual curves would contain as many as four, five or more points.

At the same time, the examination of H_2 required that at least part of the battery should be administered in a group situation. Moreover, since the theory indicated that the amount of 'guessing'[3] exhibited would be a function of

---

[3]The term 'guess' is used here, and throughout this section in a rather special sense. An individual will be said to have 'guessed' at the solution of a particular series if he offers an incorrect solution, and if it is evident from his 'stage-time' graph that he has spent an inadequate amount of time attempting the solution.

group pressure, which in turn, would depend on the overall group performance, it seemed desirable to set up a gradient of group pressure by suitably controlling the time allowance at each stage-level. Thus, if meagre time limits allowed few students to complete a particular question, the pressure exerted on the slower student to 'guess' would likely be less that if the time allowances were so ample that all except the student in question were able to complete the given series. Specifically, it was decided to control the time limits so that the number of students who would have sufficient time (at successive stage-levels) to offer a correct solution would decrease.[4] The prelininary experience with individual curves was used to estimate time limits to fit as closely as possible to the following experimental design.

Table 7

Experimental Design: The Effect of Group Pressure on Level I Performance

	Stage					
	1	2	3	4	5	7
A	100	100	100	67	33	0
B				33	67	100
C				X_1	X_2	X_3
D				Y_1	Y_2	Y_3

[4]An alternate arrangement might have been to keep the

The symbols employed in this table may be interpreted as follows:

A: Percentage of students who have sufficient time--according to their theoretical curves and the time limits imposed--to offer correct solutions at the level specified.

B: Percentage of students who have inadequate time to offer a correct solution.

C: Proportion of students in category B who 'guess', i.e., who obtain an incorrect solution in a time which is considerably less than the theoretical time required for the solution.

D: Proportion of subjects in category B who refuse to guess, and who therefore offer no solution.

If H_2 is correct, the percentage of the 'potential' guessers (B) who do, in fact, guess (C) ought to decrease as group pressure declines (A). Thus, a necessary condition for the truth of H_2 is that the X_i should form a monotone decreasing sequence.

---

percentage 'passing' at each level constant, (at 50% let us say,) and to study the guessing behaviour of the slower half. However, the experimental design employed here corresponds more closely to the normal classroom situation in which the students work on a set of exercises of increasing difficulty, with the result that the number of students completing successive problems steadily decreases.

The question of a sufficient condition for the truth of H_2 is somewhat more subtle. The membership of class C may be divided into two categories: (a) those subjects who suffer forces postulated under H_2 and (b) those subjects who are genuinely indifferent to the correctness of their answers and who consequently fill in the answers in a haphazard way merely to 'get the thing done'. Although there are several possible reasons for rejecting the second alternative, the most convincing evidence of high motivation came from first hand observation of the group in the experimental setting.

If we suppose, however, that both 'group pressure' and 'indifference' play some part in the exhibited guessing behaviour, our main concern then, becomes which factor plays the prominent part. If indifferent guessing were the more prominent cause of membership in category C, then we would expect a decrease in the 'restraint' coefficient R defined by the formula

$$R_i = \frac{Y_i}{X_i}$$

where i is the stage-level. On the other hand, an increase in R_i where i is an integer representing a step level, must be interpreted as showing the dominant influence of group pressure.

Thus, the first phase of the experimental design makes

allowances for the testing of both hypotheses, although the later arguments, removed as they are from the phenomena under investigation, do not provide the satisfaction of direct proofs. If group pressure causes upper bounds to appear, then removal from the group of an individual who shows guessing behaviour ought to result in a significant improvement in his performance--an improvement manifested in an upward extension of his upper bound. The second phase of the experiment is intended to consolidate the findings of the first phase by examining changes in the upper bounds when group pressure is removed.

2. The Test Battery. It was pointed out that the prelininary investigation showed that the lawfulness of the relationship between time and stage increased if the value of t for a given S was calculated as a mean or median t for several problems of comparable arithmetical difficulty at that S.

With this consideration in mind the following battery was prepared (Table 8):

Table 8

Distribution of Questions by Stages in Number Series Test Battery

Stages	No. of Problems[a]	Method of Timing	Allowed Time Limit/Question in Group Testing
1	20	Total time for 2 sets of 10.	15 Sec.
2	8	individual	40 Sec.
3	4	individual	115 Sec.
4	4	individual	240 Sec.
5	4	individual	300 Sec.
7	4	individual	450 Sec.

[a]The problems may be found in Appendix C.

The group battery was administered in two individual testing periods, each of approximately 45 minutes duration. For the subsequent individual testing, the first two problems at stage-levels 4, 5, and 7 were used. Three practice series were also provided at levels 1, 3 and 5 respectively.

3. Testing Procedure. While part of the battery had to be administered as a group test, calculation of k_1 and k_2 required that the times for individual solutions be recorded.

It was necessary, therefore, to devise a method whereby the student could keep an accurate record of his own times. This was accomplished by means of a large mechanical counter ordinarily used to keep scores at athletic games. An assistant had been trained previously to manipulate the dials (twisting a dial caused an increase of one in the units register), so that the device showed the time in seconds. The assistant checked this time against a stop-watch every 5 seconds--thus the time shown on the counter differed from the true time by a small fraction of a second at any given time.

The items were presented on individual sheets of a small booklet;--tearing off the top sheet exposed the first question--this was in turn removed to expose the second question, and so on.

At a given signal, the subject would begin a particular item, recording his answer (in ink) and the time shown on the clock if he was able to finish. When the predetermined time had elapsed, the order was given to stop immediately, with no further writing. Then the clock was reset and the signal given to begin the following question.

The series above stage one were presented in a random order (within each testing period) so that the subject did not know what the time limit would be. Moreover, since the subjects were not allowed to record an answer after the time

was called, it can be said that the wrong answers do not result from a last-second attempt to 'beat the clock'.

For three days prior to the experiment, the students had been given short periods of training in which they recorded the time of intervals whose length was known to the experimenter. By checking the accuracy of those times, the author was convinced that the times recorded during the experiment were correct to the nearest second.

The effect of group pressure was deliberately emphasized in this experiment. While solving the series, the student was in the normal 'concentration' position, with the head bent toward the paper. The subjects were instructed to sit erect as soon as they had finished, on the pretext that this would give the experimenter an indication when to proceed to the next problem. Thus it was quite apparent to the slower members of the group that they were, in fact, falling behind.

It is unnecessary to labour the fact that the success of an experiment of this kind depends to a large extent on the co-operation of the subjects. This was especially true with regard to the recorded times. It was pointed out to the students that their performance would not be compared to other students' (which was, of course, not true), and that the scores would not, in any event, be used to their disadvantage. The general enthusiasm of the group--stemming

from the fact that they were being allowed to participate in a 'psychological experiment'--convinced the experimenter that he was receiving full co-operation. At the end of the testing period, which consisted of two intervals of 45 minutes each, there was general reluctance to terminate the experiment.

In the individual testing, considerable care was taken to convey to the subject the idea that he could take as long as he wished to complete a question. Moreover, it was suggested that the solution would require a considerable amount of time, and that a correct solution was more valuable than a fast solution time.

Analysis of Results

1. Preliminary Analysis. The median time for correct items at each step-level was used as the subject's time at that level. If the subject scored no items correct at some level, the median time for the incorrect items was calculated and recorded for later use. Thus there was a maximum of 6 scores available for each subject. In some cases, of course, the subject would offer no solution at a particular level and there would therefore be fewer than 6 scores for him.

The scores were used to plot a graph between S and log t. Of the 106 graphs obtained, 103 exhibited a relationship

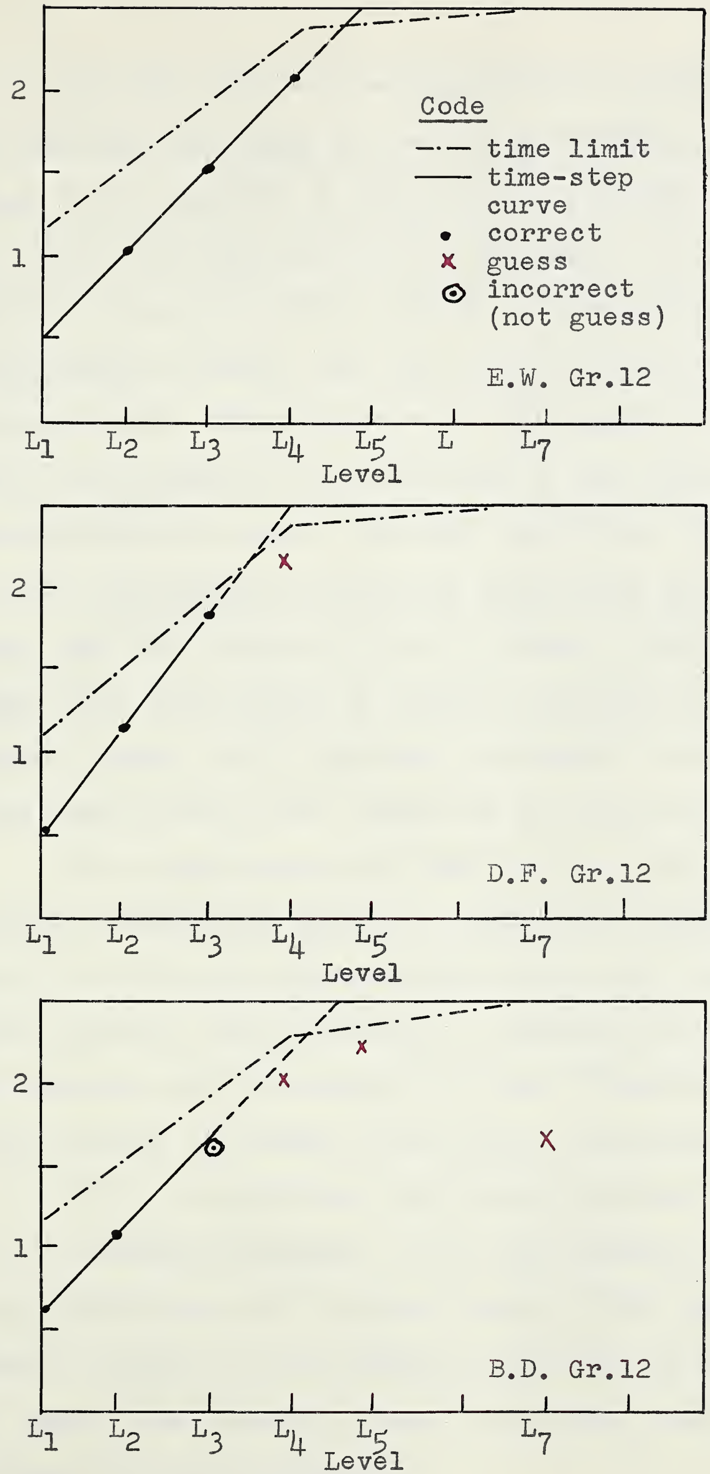
which was clearly linear over the first 5 stages. As a first analysis, each point on the graph was marked according to the following classification: (a) correct solution, (b) solution correct except for a small error which could be interpreted as a mistake in computation (e.g., 426 instead of 427), (c) incorrect solution. Three of the graphs obtained in the experiment are shown in Fig. 14.

2. The Existence of Upper Bounds. We recall that an individual's performance in a Level I-type task possesses an upper bound S_k if the subject is unable to solve all series with a number of stages S_L, where $S_L > S_k$. The distribution of upper bounds under the experimental time limits is shown in Table 9.

Table 9

Distribution of Upper Bounds in Level One Performance

	Stage					No Upper Bound Shown	Contradictory
	1	2	3	4	5		
Grade 10	7	22	19	10	5	0	4
Grade 12	0	7	11	9	6	0	3
Total	7	29	30	19	11	0	7
% (N=96)	7	30	31	20	12	0	7
Cumulative %	7	37	68	88	100	-	-



Description

Type A
No guessing-no solution offered when insufficient time given.

Type B
Guesses at level 4--but then terminates the guessing.

Type C
Continues to guess.

Fig. 14. Three Profile Types for Level-time Curves.

In order that a true upper bound S_k should exist, it was necessary that all the series containing a number of stages which exceeded S_k should not be solved. If S_{k+1} was not solved, but S_{k+t} was solved for some $t > 1$, then the subject did not, in fact, exhibit a true upper bound (in psychological sense), and the result was considered contradictory. The existence of true upper bounds in the Level I performance of 96 (of 103) of the subjects was clearly defined, and we can say, therefore, that all but 7 of the individuals solved all the series up to a certain stage, and that having failed to obtain a solution at that stage, they then failed to obtain a solution at any subsequent stage. It is in order to enquire to what extent these upper bounds were caused by the time limits imposed.

This question may be dealt with in two ways. An indirect method will allow, in addition, a check on the extent to which the experimental design was realized. Table 10 shows the percentage of subjects at each level who had adequate time, according to their theoretical curve, to offer correct solutions at the level specified.

It is evident that the desired gradient was realized to a satisfactory degree. It is also apparent that the time limits were not the major cause of the upper bounds shown in Table 9. For example, while 97% of the subjects had ample time to solve stage 3 problems, only 63% offered correct solutions (Table 9).

Table 10

Percentage of Subjects with Sufficient Time for Solution at each Stage

	Stage					
	1	2	3	4	5	7
Grade 10	100	100	97	67	29	0
Grade 12	100	100	97	75	33	0
Total Group	100	100	97	69	31	0

The problem may be dealt with more directly by referring back to the three graphs considered earlier (Fig. 14). The stage-time curves tended to be divided among these basic types as shown in Table 11.

Table 11

The Relative Proportion of Profile Types in the Individual and Combined Grades

	Type A	Type B	Type C	Not Clearly Defined
Grade 10	20	22	20	5
Grade 12	14	6	15	1
Total Group	34	28	35	6
Percentage N=103	33	27	34	6

The upper bound in the Type A profile was clearly caused by the time limits imposed. In the Type C graph, on the other hand, the subject had obtained an incorrect solution before the allotted time had elapsed, and the divergence of the solution times from the theoretical required times must be explained by some factor or factors other than the time limit. The Type B profile shows both influences, but the upper bound was set initially not by the time limit, but by guessing behaviour, so that these cases can also be regarded as showing artificial upper bounds not caused by the time limit. It is clear, then, that while restrictions in time naturally tend to cause limitations in Level I performance, these restrictions, in many classroom situations, may be of less importance than other forces.

3. Guessing Behaviour in Group Situations. Up to this point we have been prodding the periphery of the 'guessing' behaviour problem. The necessary and sufficient conditions for the truth of H_2 were previously stated as follows:

1. Necessary: the $\{X_i\}$ form a decreasing sequence,
2. Sufficient: the $\{R_i\}$ form an increasing sequence.

In order to determine the membership in class C, an exact criterion had to be established for the meaning of the term 'guessing'. Many of the incorrect answers had exhausted the necessary time as calculated from the 'step-

time' curve and could be interpreted as resulting from temporary lapses into ineffectiveness, rather than from 'guessing' in the ordinary sense of the word. The criterion established was that the point representing a 'guess' should be sufficiently below the stage-time curve that it could not be considered as a chance deviation from this curve. If the number of points was large and if we assumed 'errors', in the form of deviations, to be normally distributed about the stage-time line, then a plausible criterion would be that the deviation of the 'guess' should be of such magnitude that it could only arise 5 times in 100 by chance (i.e., that the deviation should be at least 1.96 times the standard deviations of the deviations). While the number of points in our case is extremely limited, the criterion has a definiteness which makes it useful. For most of the graphs--as for example, in our illustrative cases--inspection revealed which points represented guesses, so that the actual computation of standard deviations was required only in a few doubtful cases.

Table 12 shows the results of the computations for $\left\{X_i\right\}$ and $\left\{R_i\right\}$.

It is clear that as the gradient of group pressure declines, the proportion of those students who have insufficient time for solution and who guess, also declines. At the same time, the proportion of students in category B who

Table 12

Guessing Under a Gradient of Group Pressure

	Stage					
	1	2	3	4	5	7
A	100	100	97	69	31	0
B	0	0	3	31	69	100
X_i			100	90	66	40
Y_i			0	.10	.34	.64
R_i			0	.11	.52	1.60

show restraint in not guessing increases with the result that the coefficient of relative constraint R_i increases as well.

Moreover, the decreases in the $\{X_i\}$ are of such magnitude that differences in the proportions $D_{x_n - x_{n+1}}$ and $D_{R_n - R_{n+1}}$ are significant at the .05 level, except for the pair $D_{x_3 - x_4}$ and $D_{R_3 - R_4}$, when the small number of guesses at level three (N=3) fairly well precluded the possibility of establishing significance. In any case both the necessary and sufficient conditions for H_2 seemed to be satisfied.

So far, the guessing behaviour of a special sub-group (B) of the whole test group has been examined. This choice was one of convenience, for it provided a well-defined group with which to work. However, some of the

subjects who guessed did not belong to category B, so that it is well to consider the guessing behaviour of the whole group.

Table 13

Total Number of Students Guessing at each Stage-Level

	2	3	4	5	7
Number 'Guessing'	2	12	37	54	41
Percentage Guessing Who Are Below the Group Median	100	100	80	72	53

The table indicates that the subjects who guessed at the lower levels were those who fell behind the group. At the higher stage-levels, however, some subjects of superior performance (relative to the group) exhibited guessing behaviour.

The subjects in the latter group were often those who had fast initial times, and who were able to keep well ahead of the median group performance at the lower stage-levels. From interviews conducted in connection with the individual testing, it was learned that these subjects develop personal expectations in terms of their performance relative to the group. For example, the subject who had by far the fastest times at all levels up to and including stage 4, began to guess at stage 5, although he still had adequate time for

solution. When questioned about his faulty solutions at stages 5 and 7, he explained that his experience with the group had shown him that he usually could solve mental problems as fast as anyone in the group. In the series experiment, he found--true to his expectations--that he was the first to be finished at the lower levels, but while working on a stage 5 problem, he noticed that some students had already finished (he did not realize, of course, that these students had guessed). He stated that he was temporarily flustered, and that he began to suspect that he was "making the problem harder than it really was".

Other students of superior performance suggested that they had adjusted their performance to the speed of two or three rivals. The slower subjects almost invariably expressed their embarrassment at falling behind the group. Many of them admitted that they had guessed at the higher levels, protesting that they did not wish to 'look stupid'.

In summary, it was very clear that the performance of the group of subjects was influenced by pressures emanating from within the group. The first source of group influence was manifested as a fear of falling noticeably behind. For those subjects whose superior performance abnegated this fear, a second source of pressure was exerted via an expectation concerning their performance relative to a certain segment of the group. While the first source

seemed to be the most powerful, both were capable of causing guessing behaviour.

4. <u>A Closer Study of Guessing</u>. The word 'guessing' has been frequently used throughout this section to mean the inadequate treatment of a series resulting from an under-expenditure of time. An examination of the answers, however, indicated that in this, as well as in other tests to be considered subsequently, little genuine 'guessing', in the ordinary sense of the word, actually occurred. To give concreteness to the discussion, the diagram for the example considered on p. 103 is reconstructed, together with a diagram representing a 'guess'.

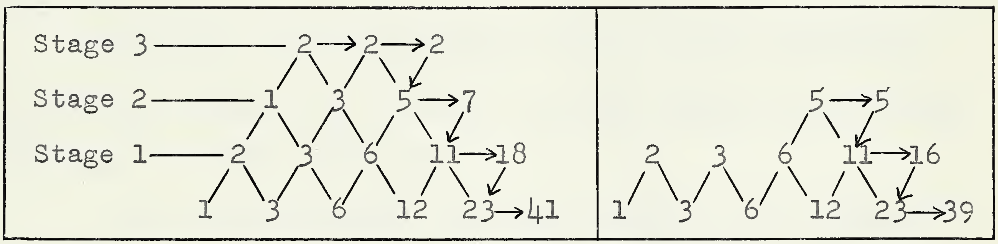


Fig. 15. A schematic representation of 'guessing' in number series.

The guess in this illustration resulted from the simplifying assumption that the series is a two-stage, when in fact, it is a three-stage series. The assumption resulted from taking inadequate time, either to store the first row of reference points or to operate on this storage. Subjects

who offered the above solution for an S_3 series, usually treated all subsequent series as though they too were S_3's.

In other words, a student who was under pressure to guess usually made what to him seemed to be a reasonable simplification, by ignoring part of the data.

5. Individual Performances. The battery for individual testing consisted of two problems at each of the stage-levels 4, 5, and 7.[5] Under the conditions of individual testing, the subject was allowed and encouraged to take all the time he needed; the result was a general raising of the upper bounds.

We first consider two of the profiles presented earlier.

We notice that the subject 'B.D.', while extending his correct answers up to level 5, still exhibited guessing behaviour at stage-level 7.

The distribution of the upper bounds with unlimited time allowance is shown in Table 14. In order to compare performance in the individual and group testing situations, the time limits for the latter situation were superimposed

---

[5] It had originally been intended to employ both grades in the individual testing. However, the extremely long intervals of time required for stage-level 7 made this impracticable; consequently, only the grade 12 group was used.

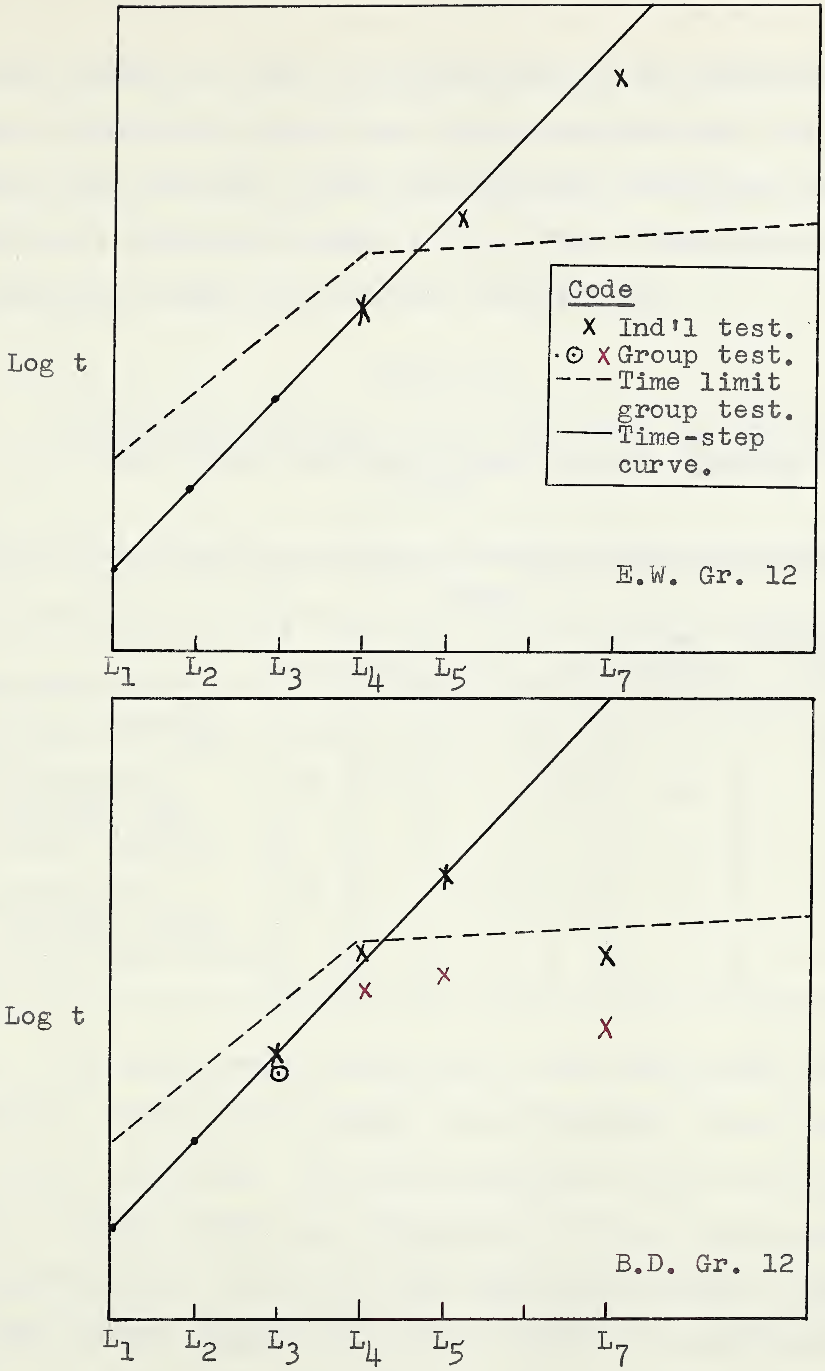


Fig. 16. Two profiles with individual and group results in Level One Performance.

on the former, so that it became possible to calculate the upper bounds which would have been exhibited had the time limits been imposed. Thus both subjects considered above would have had upper bounds at S_4. The distribution of these upper bounds also appears in Table 14.

Table 14

Distribution of Upper Bounds in Level One Ability
Under Three Different Conditions of Testing
(N=36)

	Stage						
	1	2	3	4	5	7 or above	Contradictory
A. Group Testing	0	7	11	9	6	-	3
B. Individual testing (no time limits)	0	0	0	7	12	17	0
C. Individual testing (time limits as in A).	0	0	3	23	10	-	0

The upper bounds exhibited in the individual testing situation seemed to be higher (significantly) than those in the group situation. The null hypothesis that the distribution of upper bounds was independent of the testing situation was tested (for A and C) by the sign test (Siegel, 1956, p. 68) and was rejected at the .01 level of significance.

A substantial part of the raising of the upper bounds

resulted from a decrease in guessing, particularly at the lower levels.[6] There was also a decrease--again at the lower levels--of those points which represent incorrect answers, and yet which result from a complete consumption of the available time (these were earlier attributed to temporary lapses in efficiency).

The distribution of finite upper bounds under category B is somewhat disturbing, for from the point of view of the theory, no natural upper bounds ought to appear if the student is given sufficient time. However, while guessing behaviour is reduced in this face-to-face testing situation, it is by no means eliminated. Many students seem to carry their performance expectations (in terms of time) into the individual testing situation, so that, they manifest guessing behaviour even under the strongest assurances from the experimenter. Again, it is likely that some students would be more self-conscious concerning their slow performance in face-to-face situations, than if they are working in the group. Indeed, the author would not be tempted to conclude that guessing effects tend to be less in

---

[6]Since the times obtained in situation C continued to provide a good fit to the 'stage-time' curve obtained in situation A, the effects of practice could not have been large enough to contribute substantially to the general raising of the upper bounds. Actually the theoretical curve gave a close fit to the observed values for the first 5 levels; at level 7, however, the theoretical curve tended to overestimate the observed values.

'face-to-face' situations in general--an experiment reported later shows that momentary 'sets' can be induced in the face-to-face situation which result in extreme displays of guessing. What we can conclude is that in this particular experiment, it was possible to make significant reductions in the amount of guessing manifested by the subjects.

Further Statistical Considerations

1. Reliability Estimates. The fact that the points at successive levels fit closely to a straight line, itself offers some evidence for the reliability of the series at each level. However, a more direct calculation was possible, because the items--at least above stage one--were timed individually. The reliability coefficient at stage-level one was determined by calculating a product-moment coefficient of correlation between the average time per series, determined by dividing the total time of each sub-test (10 questions) by 10. At stage-level two, the median time for the odd-numbered series was correlated with the median time for the even-numbered series. Beyond S_2, the frequency of errors and the occurrence of upper bounds made the calculations impossible. However, it is probably safe to assume that reliability coefficients of the same order would exist at these upper levels.

Table 15

Reliabilities for One-Stage and Two-Stage Problems in Level One Number Series

		Grade 10 (N=67)	Grade 12 (N=36)
S_1	Split-half	.70	.74
	Corrected by Spearman-Brown	.83	.85
S_2	Split-half	.73	.76
	Corrected by Spearman-Brown	.85	.87

The reliability of the whole battery at both levels was estimated by means of the Spearman-Brown formula (Garrett, 1953, p. 341).

2. <u>The Constants k_1 and k_2</u>. The 'time-step' equation can be written in the form

$$\log t = k_1 + k_2 S$$

Mathematically, k_1 is the intercept of the straight line on the log t axis--it represents the time for one-stage problems corrected by means of the general configuration of the points constituting the straight line. Since the line often passed through the initial point, a very high relationship between L_1 and k_1 would be expected, so that the product moment coefficient of correlation of .93 is not surprising. Moreover, since no direct estimate of the

reliability of k_1 is available, we will assume that k_1 has the same degree of reliability as the closely related L_1.

The constant k_2 is the slope of the stage-time curve, and represents the rate at which log t increases from level to level. Here, no direct estimate of the reliability of k_2 was possible.

The values of k_1 and k_2 were computed from the graph for each subject, and a further investigation was made concerning the psychological meaning of these constants. Table 16 presents means and standard deviations of k_1 and k_2 for the matched groups. (The results for the whole groups were comparable.)

Table 16

Means and Standard Deviations of k_1 and k_2 for the Matched Groups

	Grade 10	Grade 12	Total
Mean k_1	6.41	6.85	6.63
S.D. k_1	2.28	2.78	2.59
Mean k_2	46.5	47.5	47.0
S.D. k_2	9.60	10.9	10.3

Table 17

Correlations Between k_1, k_2 and Otis for the Matched Groups

Grade 10		Grade 12		Total	
k_1	k_2	k_1	k_2	k_1	k_2
-.63	.35	-.63	.33	-.63	.33

The constant k_1 represents the speed of performing subtractive mental operations with little storage, and undoubtedly depends upon the hypothesized attention-cathexis or mental energy which the subject has at his disposal. This ability does not seem to change much between the grade 10 and grade 12 levels; in fact, the data show a slight advantage (not significant) in favour of the grade 10 group. This would suggest that once the stage of formal mental operations is reached, the 'peak' attention-cathexis level tends to stabilize, so that advances in mental performance result primarily from improved strategy (i.e., advantageous deployment of existing hypercathexis) rather than from an increased speed of mental operation (increased hypercathexis).

Since the speed of performance for simple mental operations would be proportional to the available hypercathexis (or mental energy), it is not surprising that a fairly high correlation existed between k_1 and Otis IQ.

It is probable that this coefficient is contaminated somewhat due to inequalities in arithmetical training and experience. In a later experiment a 'pure' mental operation (i.e., not influenced by previous experience), yielded coefficients with the Otis in the neighbourhood of .80.

The constant k_2, which again showed no significant change from the grade 10 to the grade 12 groups, probably should be interpreted as representing a storage function. The barely significant (.05 level) coefficient of correlation between k_2 and the Otis suggests that the individual who scores high on the Otis performs the initial simple mental operations much more rapidly than the slower subject, with the result that there is a slight tendency for him to show a greater proportionate increase in time required for solution from step to step. However, since k_2, in contrast to k_1, lacks an independent existence, arguments employing this constant can only claim second-order plausibility.

Summary

Two main results emerge from this investigation. In the first place, it is clear that the formula provided a workable estimate of the comprehension time for the number series constructed by the calculus of finite differences. Graphs obtained in group and individual testing situations corroborated this finding. Not only does this have

practical value in that it provides a measure against which group-induced perturbations in a special type of cognitive behaviour can be studied, but it has theoretical value in that the equation constants (k_1 and k_2) can be given further psychological meaning.

The constant k_1, representing the speed at which a simple type of mental operation is performed, correlated highly with the Otis IQ. The energy theory proposed in the earlier sections would suggest that the speed of mental operations depends upon the amount of attention-cathexis available--so that the observed high correlation indicates that Otis IQ is substantially a measure of the average quantity of attention-cathexis which the subject brings to bear on the mental task. Under 'steady state' conditions--high motivation, and short intervals of time--mental performance was fairly constant, yielding reliability coefficients of the order of .85.

The constant k_2 suffers in that it is secondarily derived from the level-time curves, and so one must be cautious in its interpretation. However, it seems to represent a storage function, and if we sought a physiological correlate, we should be tempted to think of the natural resistance in the neural circuits which Hebb conceptualized as 'phase sequences'. In any case k_2 showed a small, but positive correlation with the Otis, indicating that high IQ

tends to be reflected in a type of mental organization which operates at high speeds at the lower levels but requires relatively large increments of storage time in passing from level to level.

The second finding was that Level I performance exhibited many instances of induced upper bounds in a group testing situation structured to approximate normal classroom procedure in which students attempt a set of exercises of increasing complexity. In this situation, the amount of 'guessing'--which is manifested in a more or less systematic ignoring of part of the given data--covaried with a gradient of group pressure defined by the proportion of students in the group who were able to obtain solutions in the time allotted.

Part of the test was repeated in an individual testing situation in which the subject was encouraged to use sufficient time to obtain a correct solution. The fact that this led to significant increases in the upper bounds under the same time limits gave further support to the hypothesis that these upper bounds are, to some extent, group-induced. Further inquiry led to the belief that the subject tended to evaluate his own performance (and to adjust his solution times) in terms of the performance of the whole group or some segment of it.

The preceding example is part of a larger thesis which considers the operation of so-called 'logical' behaviour--i.e., the unbiased examination of the entire body of relevant evidence--to be very much dependent upon the situation. Indeed, the student in the classroom who finds guessing less painful than complete thinking is not in an unusual situation, for in everyday life we normally exhibit just sufficient logic to disguise the incompleteness of our considerations, and we are easily forced into illogical behaviour through the desire to avoid unpleasant situations.

In any case, this virtually unexplored area is of great importance to anyone who would consider group performance or testing, and it is returned to later in a different context.

CHAPTER IX

GAP FILLING AND STRATEGIES

The theory predicts that the occurrence of a gap in a deductive sequence will cause an increase in comprehension time, because the individual is obliged to search among the possible operations of the system in order to locate the one which has been applied. Although a purely random search is possible in theory, two factors militate against its occurrence in practice: (1) The immediate context provides clues as to which operation has been applied; (2) The recurrence of a certain type of gap leads to the development and application of a strategy for dealing with it.

The results of a previous experiment suggested that once the human brain attains the stage of formal operations and has available a quantity of hypercathectic energy which fluctuates and yet which possesses a definite upper bound--the maximum rate of performance of simple mental operations does not increase. At the same time, some aspects of mental performance, as shown on intelligence tests for example, reach a maximum about the twentieth year of age. There is here, an implication that the increase in level of performance results not so much from an increase in the quantity of attention-cathexis, but rather from an increased efficiency in its expenditure. This takes place through the

development of strategies which formulate systematic methods of approach to situations which tend to recur in the subject's experience.

Three hypotheses thus emerged to give direction to the experimental procedures:

1. H_1: Random, or near-random behaviour--identified by the absence of a strategy--leads to increases in comprehension time which becomes greater as the gap lengthens.

2. H_2: The repetition of a certain type of gap gives rise to the development of strategies for dealing with it.

3. H_3: The gap-filling behaviour of a grade 12 group is characterized by strategy formation to a greater extent than a grade 10 group equated on Otis IQ.

1. <u>The Test</u>. In order to provide a simplified gap situation, a test was devised involving permutations on four elements.

A B C D

Fig. 17. Permutations on four elements.

The permissible primitive operations in this simple system are the six permutations: I_{12}, I_{13}, I_{14}, I_{23}, I_{24}, I_{34}. The letter 'I' may be translated as 'interchange' and the subscripts indicate the position of the elements to be interchanged, so that the operation I_{24} performed on the original 'axiom' yields the result

A B C D

Second and third-order operations are defined in a similar

manner; thus, the necessary mental processes involved in the comprehension of the operation (or sequence of operations) $I_{34,\ 24,\ 12}$ may be elaborated as follows:

Initial Reference Point	A	B	C	D	
I_{34}	A	B	D	C	Storage 1
Operation on Storage - I_{24}	A	C	D	B	Storage 2
Operation on Storage - I_{12}	C	A	D	B	Result

Fig. 18. Sequence of mental steps in the solution of a third order permutation.

The performance required here is the typical Level I sequence of: operation$\longrightarrow$storage$\longrightarrow$operation-on-storage. The permutation operation is perhaps, a 'purer' one than the arithmetical operation used previously, in that it was unlikely that any of the students had had extensive practice with it.

So far, only the 'comprehension' side of the test has been considered. Suppose, however, that the subject is presented with the original 'axiom' and result, as illustrated in Fig. 19. He is, moreover, asked to find the two operations which have been performed on the axiom to produce the result. It is conceivable that he might first consider each of the 6 permissible operations which could

have been performed on the axiom, store each result, and then consider each of the 6 possible operations applied to each of the 6 stored results. In this way, he could pick out the pair or pairs of operations which combined to give the answer.

Axiom:	A	B	C	D	
Operation 1					Storage 1
Operation 2	C	A	B	D	Result

Fig. 19. Gap filling in the permutation sequence.

However, some modification of this completely random behaviour is invariably in evidence. The subject may first notice that the element C is out of place and may perform operation I_{13} on the original set, thereby moving C to its correct position in the result. The element A would now occupy the third position and the further operation I_{23} is necessary. This is one of the possible strategies considered in the later analysis.

2. <u>Test Battery</u>. The test battery was designed to give a representative measure of both comprehension and gap-filling times for each of the first, second, and third order operations.

The comprehension questions provided the operations, while the gap-filling questions provided the results and

inquired about the operations; thus the questions were in one-to-one correspondence. The battery included 20 of the 24 possible permutations on the four letters A, B, C, D--the four excluded were the identity permutation (a simple first order operation) and three second order operations which did not involve moving at least one element twice and which had been found to be unrepresentative of that level. Since the multiple operations may be stated in more than one form, care was exercised to ensure that no particular element, position, or anticipated strategy would be favoured.

Table 18

Permutation Battery For Gap-Filling Test[a]

	Comprehension	Gap-Filling
1st Order	6	6
2nd Order	8	8
3rd Order	6	6

[a]The individual problems may be found in Appendix C.

3. <u>Subjects and Testing Procedure</u>. The 60 students of the matched groups were tested individually and the solutions were timed with a stop-watch. As a preliminary introduction to the notion of permutations, the simpler set of three elements was considered, but no strategy was

suggested.

The comprehension questions at each level were given first, and were immediately followed by the gap-filling questions at the same level. The entire test required approximately 45 minutes of administration time per student. No computing aids of any kind were allowed for either test.

At the mid-point and end of each level of the gap-filling tests, the subject was interrogated to determine which method he was using to fill the gaps. Thus 6 such check points were established, so that the experimenter was able to use these verbal reports, in conjunction with the answers themselves, to determine for each subject, the degree and type of strategy formation and the consistency of application employed at each level.

Analysis of Results

1. Comprehension Scores. The subject's score for each level and on each test was determined by computing the median time (in seconds) for the correct solutions offered. Table 19 presents some data concerning the comprehension scores.

The difference between grade 10 and grade 12 means was not significant at any level, providing further evidence that the speed of performance of simple mental operations

does not change radically after the stage of formal operation is reached.

Table 19

Means (time in seconds) and Standard Deviations for L_1, L_2, L_3 Scores for Matched Grade 10 and Grade 12 Groups

		L_1	L_2	L_3
Grade 10	M	5.85	13.1	25.7
	S.D.	1.58	5.60	11.2
Grade 12	M.	5.71	12.3	24.3
	S.D.	1.28	5.78	11.4
Total	M	5.78	12.7	25.0
	S.D.	1.46	5.68	11.3

The data provided some corroborating evidence for observations made in the previous experiment. We first notice that the increases in the mean group times from level to level are again nearly exponential--the time roughly doubling between levels. A majority of individual profiles show a fairly good fit to an exponential curve, but since only three points are available, the relationship was somewhat tenuous and the curves were not drawn.

An odds-evens product-moment reliability coefficient was determined for the 60 scores at L_1 and amounted to .77, with a corrected (Spearman-Brown) coefficient of .87.

The interrelationship of the L_1, k_1 and Otis scores are shown in the following table.

Table 20

Product Moment Correlations Between the k_1, L_1 and Otis Scores

	Grade 10		Grade 12		Total	
	L_1	Otis	L_1	Otis	L_1	Otis
k_1	+.87	-.63	+.83	-.63	+.84	-.63
L_1		-.79		-.80		-.80

The two matched population were pooled on the grounds that they are comparable in mean and standard deviation and the scattergram reveals one homogeneous cluster of points rather than two separate clusters.

The L_1 scores correlate somewhat higher (-.80) with the Otis than does k_1. It was suggested previously that the 'pure' permutation operation would probably be less contaminated through previous experience. The k_1, L_1 correlations are very high, exceeding, in fact, the correlations of the Otis with either test. In this connection, one is almost tempted to postulate the existence of a mental factor--an attention-cathexis powered facility in the performance of simple mental operations.

2. Gap Filling. Table 21 presents the essential data concerning the gap-filling scores at each level and for each grade.

Table 21

Means and Standard Deviations of the G_1, G_2, G_3 Scores for Matched Grade 10 and Grade 12 Groups

		G_1	G_2	G_3
Grade 10	M	5.48	30.6	71.9
	S.D.	2.11		
Grade 12	M	5.39	24.0	43.8
	S.D.	2.12		
Total	M	5.34		
	S.D.	2.12		

We notice first the departure from logarithmic form; the increment in times between G_1 and G_2 is nearly quadrupled (doubled for comprehension times) while the times are doubled (approximately) between G_2 and G_3. There is an indication here of the development of time-saving strategies between G_2 and G_3. Moreover, the significant differences between grade 10 and grade 12 scores at G_2 and G_3, coupled with the non-significant differences in comprehension times, suggests that the grade 12 group was, in fact, exhibiting a greater amount of strategy formation. This provided an

indirect confirmation of H_2. We also notice that while the gap-filling time at G_1 was slightly less than the comprehension times, the gap-filling times for G_2, G_3 were markedly and significantly larger than the comprehension times. These microscopic considerations will be followed by closer observations of individual strategy formation.

Strategy Analysis

We recall here that a strategy has been defined as a non-random method of dealing with recurring situations within a certain area of experience. The essential requirements of a strategy are consistency, explicit definition, and effectiveness, and the overall 'goodness' of a strategy is measured by the gain which it affords in time or effort over a purely random approach.

Within the experimental framework, three types of strategy seemed to emerge, and are discussed briefly in order of their frequency of occurrence.

Given Axiom	A	B	C	D	
Operation 1					Storage 1
Operation 2					Storage 2
Operation 3	C	A	D	B	Result

Fig. 20. Schematic representation of a permutation gap of length three.

(a) Horizontal Position Correction. By far the greatest proportion of strategies (approximately 75%) proceeded by performing operations in a left-to-right direction which moved the out-of-order elements into their new positions. Thus, C is the first element in the result which is out of order; consequently the operation I_{13} must have been performed on the original set, resulting in the new order

C B A D --Storage 1

A is the next element whose position in Storage 1 differs from its position in the result; thus the next operation, performed this time on Storage 1, must have been I_{23}. The new Storage is,

C A B D --Storage 2

Finally, the element B is out of position, and the required operation on Storage 2 is I_{34}. Thus the required sequence of operations is $I_{13,\ 23,\ 34}$. The high frequency of strategies which followed the normal direction of movement in reading and writing is probably to be expected.

(b) Alphabetical Position Correction. Approximately 25% of the strategies employed a procedure in which the positions of the elements are corrected in order of their occurrence in the alphabet. Thus, the position of A is corrected first, by performing the operation I_{12}, on the original set, yielding the order

B A C D --Storage 1

The position of (B) is corrected next, the operation I_{14} on Storage 1 yielding

(D) (A) (C) (B) --Storage 2

The final operation I_{13} on Storage 2 corrects the position of (C) and the required sequence is $I_{12, 14, 13}$.

This solution yields a different answer from the horizontal-position-correction strategy; there are, in fact, various answers, the number of which depends on the level.

Table 22

The Number of Possible and Correct Answers in Relation to Gap Length

Level	Correct Answers	No. of Possible Answers
G_1	1	6
G_2	3	36
G_3	14	216

It is apparent that the probability of obtaining a correct answer by chance is so small that it can be ignored.

(c) Single Position Correction. The most efficient strategy greatly reduced the memory load in the solution, and involves a modification of the preceding strategies. Here we may work from any of the four positions--we use the second for

purposes of illustration. The element A occupies this position in the result; thus B has been moved out of this position, implying the performance of I_{24}. This operation would bring D into the second position--and so I_{23} is required to move it into the third position. This operation brings C into the second position and I_{21} is required to move it into the first position. Thus the sequence of operations is $I_{24,\ 23,\ 21}$. This strategy--which requires only the retention of the position of one element--was clearly formulated and consistently applied by only one student in the group of 60.

1. Criteria for Rating Strategy. Strategy formation probably varies along dimensions of increasing explicit formulation and consistent application. The 6 verbal reports, in conjunction with the answers given by the student, were used to assess the degree of strategy formation at each level. The performance at a particular level was assigned the label describing the degree of strategy formation which was in evidence over the larger part of that level. The criteria for the three levels were set down as follows:

A. Complete Strategy: This was defined to be any strategy with the following properties: (a) Complete consistency of application, (b) Complete accuracy of application.

B. Partial Strategy: This was defined as any strategy which violated either (a) or (b), but not both, once in three applications or less frequently.

C. Inadequate Strategy: This was defined to be any strategy which violated either (a) or (b) more frequently than once in three applications, or which violated both criteria to any extent at all.

The following table provides a summary of strategy formation in both grades.

Table 23

Strategy Formation at Levels G_2 and G_3 for the Matched Groups in Grades 10 and 12

	L_2				L_3			
	A	B	C	A+B	A	B	C	A+B
Grade 10	1	4	25	5	1	7	22	8
Grade 12	4	12	14	16	8	12	10	20

It is apparent that the grade 12 group is superior in categories A and B. If we combined these two categories, thereby dividing each of the matched groups into 'strategy' and 'non-strategy' subgroups, the superiority of the grade 12 group is demonstrable statistically.[1]

[1]The gap of length one was so easy to complete that the detection of a strategy and its division into categories A, B, and C was not possible.

The hypothesis that the two matched groups are drawn from the same population or populations with equal proportions of 'strategy' and 'non-strategy' members was tested by means of χ^2 and was rejected at the .05 level in both cases. Thus the superiority of the grade 12 group in strategy formation has been demonstrated directly and indirectly.

2. Gap-Filling Time Curves for Strategy and Non-strategy Groups. The mean times for each matched group and the combined group have previously been considered. It is also useful to consider the effects of the various levels of strategy formation upon the time required for solution. Fig. 21 illustrates this situation.

The results are obvious enough. A superior strategy allowed gaps to be filled in times which were less than the corresponding comprehension time; an inconsistently applied strategy required times whichwere somewhat larger than the comprehension time; a very weak strategy soon caused reconstruction times to become formidably large.

The preceding discussion has some important implications for the normal mathematical exposition. Here, the more-or-less random occurrence of types of gaps precludes the possibility of a systematic development of strategies, so that when gaps occur, they become a formidable, if not

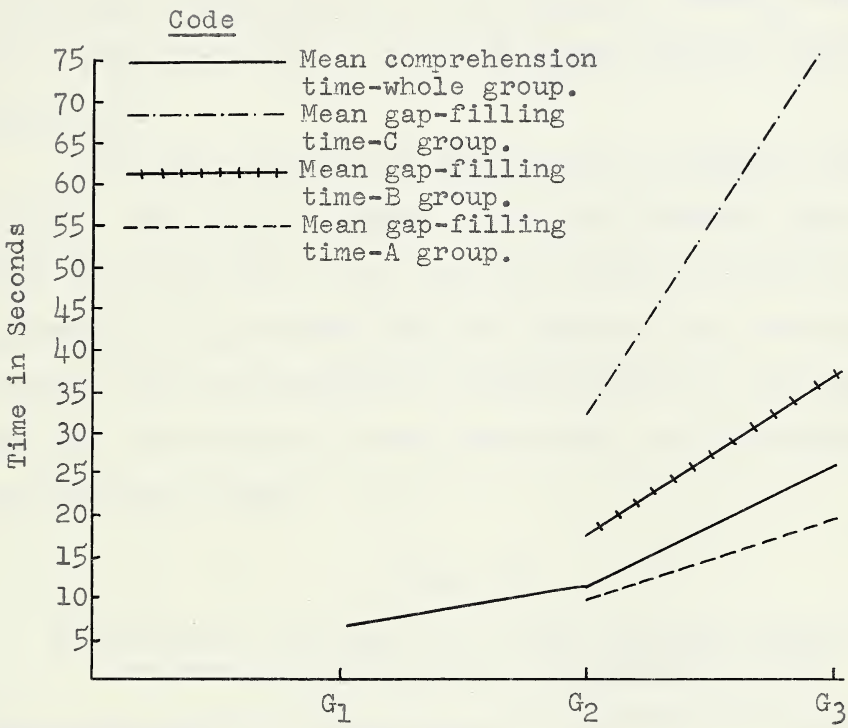


Fig. 21. Gap length and comprehension time under three levels of strategy formation.

insurmountable barrier to comprehension.

Strategy Formation and IQ

One would expect IQ and strategy formation to correlate highly. This seems to be implicit in 'adaptation' definitions of intelligence--for superior strategy formation over a wide variety of activities surely implies superior adaptation.

The data provided an opportunity to examine these beliefs. The product-moment correlation between gap-filling times (G_1) and Otis IQ for the whole group amounted to .57, indicating a significant but not markedly high relationship. To study the relationship at levels G_2 and G_3, the population was dichotomized in two dimensions, as illustrated in the following table.

Table 24

Relationship Between Otis IQ and Strategy Formation at Levels G_1 and G_2 for the Combined Groups

	Above Mean IQ	Mean IQ		Mean IQ	Below Mean IQ
Adequate Strategy	9	12	Adequate Strategy	13	15
Inadequate Strategy	21	18	Inadequate Strategy	17	15
	N=60			N=60	
	G_2			G_3	

An estimate of the relationship at each level was found by computing corresponding phi-coefficients (Garrett, p. 367). These turned out to be -.10 and -.037 at levels G_2 and G_3 respectively; consequently, we can say that the data do not bear out the hypothesized high relationship between IQ and strategy formation.

Since this result runs somewhat contrary to normal expectations, a supporting result (which is not directly related to gap-filling behaviour) is introduced here. A series of 'ordering' problems (Appendix B) of the type

T is less than L.
B is greater than A.
A is greater than L.
Which is the greatest?

gave rise to a second strategy-formation situation; again, it was possible to establish criteria for judging the effectiveness of the strategy. When the population was dichotomized, as in Table 24, and phi-coefficients calculated, the following table was obtained.

Table 25

Relationship Between Otis IQ and Strategy Formation on Ordering Tests, Grades 10 and 12

	ϕ	Significance
Grade 10, N=67	.30	$.01 < P < .05$
Grade 12, N=37	.35	$.01 < P < .05$

In this case the individual coefficients, although relatively small, were both significant at the .05 level, and the combined results were significant beyond the .01 level.[2] The relationship between IQ and strategy formation--as it emerged from the two tests employed--was somewhat complex and is represented diagrammatically in Fig. 22.

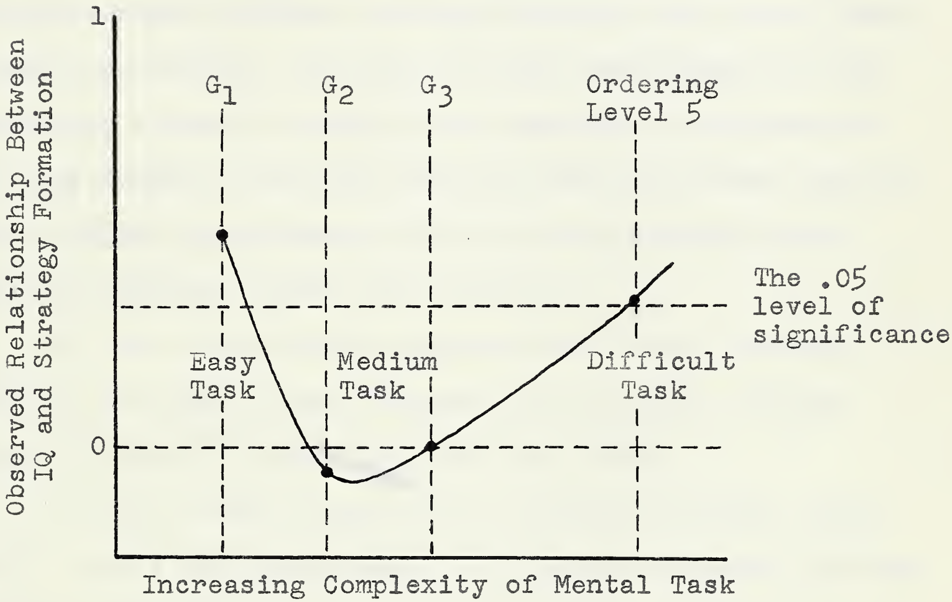


Fig. 22. The relationship between Otis IQ and strategy formation as a function of task difficulty.

[2]In addition, a comparison of the matched grade 10 and grade 12 groups employing the sign test showed the latter group to be significantly superior in strategy formation. ($.01 < P < .05$).

Several explanatory remarks are required to interpret this curve.

(1) The difficulty of obtaining an estimate of strategy formation for mental tasks of small gap length was mentioned previously. If, however, the speed of performance is taken as the best estimate of strategy formation, then the observed product-moment r of .57 gives an estimate of the relationship between strategy formation and IQ in tasks of small gap length. The size of this coefficient is not surprising, since the mental task required in problems of small gap length is not far removed from the direct application of mental operations, and this has previously been shown to correlate highly with Otis IQ.

(2) The relationship between Otis IQ and strategy formation in tasks of gap lengths two and three yielded negative phi-coefficients of near zero value.

(3) The longer tasks of the ordering battery again showed a small but significant relationship between strategy formation and IQ.

One possible explanation of the low relationship exhibited in the centre of the curve is that the more intelligent subject is able to perform simple mental operations quickly and is thus able to succeed in tasks of median difficulty while employing a strategy which would be inadequate if used by the weaker student. The observation of

students attacking the ordering problems supported this hypothesis. Many of the superior IQ students persisted in the use of an inadequate strategy (in this case memorizing the position of the elements). As the problems became harder there occurred for these subjects, a period of reorganization and experimentation, during which their exhibited strategies were inconsistent. At the same time, the low IQ student, if he developed a strategy at all, tended to do so earlier, because his relatively inferior performance of mental operations (and storage) made the early problems more difficult for him. In any case, the relationship between IQ and strategy formation was in no case very high.

Summary

This section provided further evidence for the hypothesis that the time function of the operation storage operation-on-storage sequence is exponential in nature. High intercorrelations of Otis IQ, K_1 and L_1 suggest that these three tests measure, to a large degree, the ability to perform simple mental operations, an ability which is linked in theory to general energy level (hypercathexis).

A simplified gap situation was created in which the subject had to reconstruct a sequence of omitted operations.

This setting allowed a study of strategy formation, i.e., ways of effecting savings in time through systematic modes of procedure in situations which tend to recur. While strategy formation varied along a continuum--from a near-random approach on one extreme to well-defined, consistent strategies on the other--three levels were distinguished for the purpose of experimentation. As might be expected, particularly well-defined and effective strategies resulted in a saving (as compared to comprehension time) over a gap of any length. However, poorly-defined or inconsistent strategies led to rapidly increasing increments in required solution time (above ordinary comprehension time).

Two important findings emerged concerning strategy formation. The first was that while the grade 10 and grade 12 groups showed comparable comprehension times, the latter group was significantly superior in strategy formation on two distinct tests.

A second, and rather provocative result, was that for the students investigated here, the relationship between Otis IQ and strategy formation was complex, but not markedly high. In other words, the formation of a successful strategy seemed to depend on some quality over and above IQ. Indeed, after observing such behaviour over a period of 150 hours, the author was convinced that many students of high ability persisted in the use of inadequate strategies

because their superior performance (of operations) allowed them to succeed where weaker students (using the same strategy) would fail. Often, then, the strategy formation of the subject of mediocre ability resulted from a necessity not felt by the subject of high IQ.[3]

The concept of strategy[4]--although virtually neglected (outside the theory of games) at the present time--may ultimately provide a useful supplement to our speed-oriented concept and tests of intelligence. There are long range aspects of adaptation which would be universally accepted as manifestations of intelligence, and yet which may not be fairly assessed by the standard, short interval intelligence tests.

---

[3]This discussion brings to mind the related and apparently unresolved controversy as to whether Einstellung is more prominent in intelligent or unintelligent subjects. (Miller, 1957).

[4]The repeated application of tests of cognitive ability gives rise to variability in scores which is often discussed in terms of 'function fluctuation' (as distinguished from test error). One determinant of fluctuation must surely be linked to the differential development of strategies.

CHAPTER X

THE COMPREHENSION OF TOPOLOGICAL CONCEPTS

The movement on foot for the introduction of 'modern' mathematics into the high school accepts, although seldom states explicitly, two assumptions: (a) it is possible to teach modern mathematics so that it will be comprehended by the high school student, (b) there is something to be gained in doing this.

The second assumption may never be more than mere opinion, and the first has received scant attention. The account which follows describes an experiment in which grade 12 students were introduced to general topology. This occasion offered an opportunity to study the effect on comprehension of the lecture method of mathematical exposition, and to investigate the difficulties encountered by students in the comprehension of axiomatic mathematics.

Subjects and Procedure

The grade 12 group was available for this experiment since their course of study was somewhat more flexible than that of grade 10. This group was divided into two sections, subsequently referred to as Group A and Group B. The groups were equated in intelligence (mean Otis IQ of 111) and were of roughly the same general accomplishment in their previous

work in mathematics. The experimenter was able to do this by virtue of his acquaintance with, and participation in, their previous mathematical training. Augmented by three grade 11 students, the groups numbered 20 students each.

The material to be covered was taken from the first few pages of Sierpinski's General Topology. It represented approximately the amount that would be covered in a 50 minute university mathematics lecture. The subject matter headings in their order of presentation are listed as follows:

1. Undefined terms: element, set, subset, null set.
2. Definition: Frechet (V) space.
3. Definition: limit point of a (V) space.
4. Definition: derived set.
5. Definition: Topological equivalence of (V) space.
6. Theorem 1: Two (V) spaces K_1 and K_2 consisting of the same elements are topologically equivalent, (we assume that each element is contained in each one of its neighbourhoods) if and only if, to every neighbourhood U, of an element in K_1 there exists a neighbourhood of that element in K_2 which is contained in U, and vice versa.

The objective in Group B was to adapt the pace and method of teaching to ensure that a certain minimum percentage of the subjects would comprehend the material. The material was delivered in daily periods of one-half hour's duration, although the full time was not always used. At the end of each numbered section, the subjects were asked

to indicate whether they felt that they comprehended the section. If 75% or more of the class responded in the affirmative, a comprehension test was given. If less than this proportion responded in the affirmative, the section was explained again and further examples were given.

Each student had been previously given a 6-page booklet--after each section had been demonstrated, the subject was asked: (a) to indicate whether he felt that he had comprehended the section, (b) to answer a question designed to measure this comprehension. The comprehension questions may be found in Appendix C.

Various means were employed to keep motivation high in both groups throughout the experiment. For example, it was intimated to both groups that the topology score would contribute a substantial share to the term mark. Thus, while estimates of group motivation are fairly well confined to subjective impressions derived from the testing situation and the students' response, it will be maintained that general motivational level in each group was high--at least at the beginning of the session.

Throughout the session with Group B, questions from the class were not only encouraged, but responses were evoked from the class in the normal teaching manner. The schedule for Group B is shown in Table 26.

The experimenter kept a close record of the time

Table 26

Lesson Schedule for Instruction in the Comprehension of Topological Concepts

Lesson	Topic Number	Content
I	1	Sets, subsets, null set
	2	Frechet (V) space
II	3	Limit point of a (V) space
	4	Derived set, closed set
III	5	Topological equivalence
IV	5	Topological equivalence
	6	Theorem I
V	6	Theorem I
VI	6	Theorem I
VII	6	Theorem I

actually spent on each section of the work by marking times on a prepared set of lesson notes which were used with both sections.

The material was delivered to Group A in a manner designed to approximate closely to the mathematics lecture. The material was set out in logical fashion, with every attempt made to reduce gaps. At each step, representatives were introduced in the form of examples (a minimum of two at each numbered step), and the development of the arguments proceeded at a very slow pace, requiring in excess of 87 minutes.

The amount of material covered per lecture was the same as the Group B, except that Theorem I was divided into two parts. Again, the 'lecturing' time was considerably less than the teaching time. Students in Group A were not encouraged to ask questions, but explanations and examples were given where they were requested. A check on comprehension time was maintained throughout the experimental period by means of the test employed with Group B.

As with Group B, no note-taking was allowed or assignments made during the experimental period. In short, the chief difference between the two experimental groups was in the method of delivering the material and in the time employed.

Results

The demonstration times for each section of the work (topic) with Group A were marked off cumulatively along a time axis (Fig. 23). Thus, in this scale of unequal time intervals, the logical-demonstration-time curve appears as a straight line. The comprehension curve, giving the proportion of students comprehending the material up to the end of a specific position (as determined from the comprehension tests) has been superimposed on the curve.

It is contended that the curve represents what happens when the student is faced with the traditional mathematics lecture. In fact, the 87 minute session in its entirety might be considered equivalent in many ways to a single mathematics lecture.

In Group B, demonstration time had been adjusted so that a relatively constant proportion of the subjects comprehended the material, thus allowing a study of the nature of the 'time-comprehension' curve. It is apparent from Fig. 23 that comprehension-time is not a linear function of demonstration units[1]--but appears rather to be an exponential type function.

From this microscopic point of view, the real problem

---

[1]The units of demonstration time were those employed in the lecture development of the material with Group A.

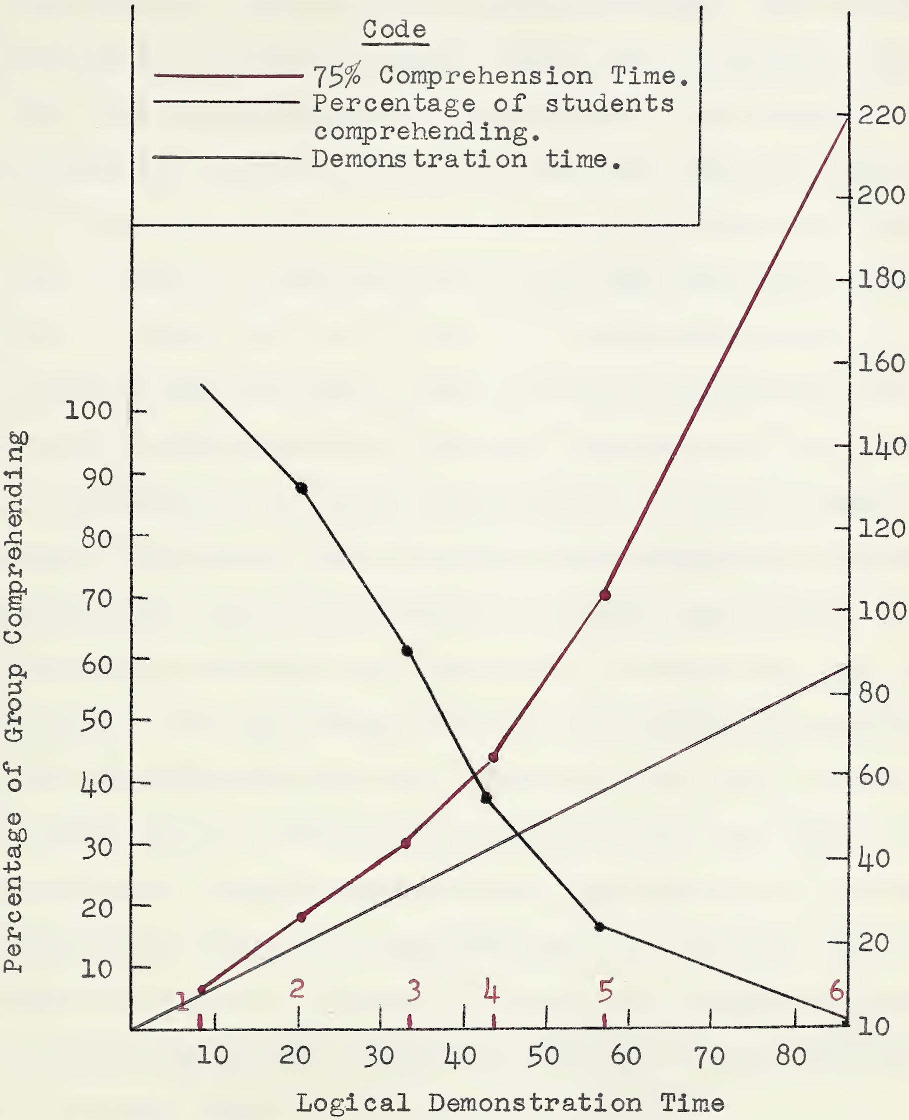


Fig. 23. Seventy-five per cent comprehension curves, percentage comprehension, and logical demonstration times for elementary topological concepts.

of linear presentation time as opposed to logarithmic comprehension time readily becomes apparent. The 'lecture' delivered to Group A, started from first principles and with 100% comprehension by the students. The normal mathematics lecture, beginning from some advanced point and utilizing the results previously acquired, usually starts from a level of comprehension below 100% and would likely have a higher rate of falling off in the proportion of students comprehending. Under these circumstances, the explicit advantages of a 'lecture' approach to the learning of mathematics is highly questionable. It is, in fact, almost traditional that students find mathematical expositions, even those which purport to begin from first principles, difficult or impossible to follow for any length of time. It would seem a tenable conclusion, that mathematical comprehension must be 'taught' in the sense that a contact must be established with the group, and that adjustments in demonstration rate must be made to fit the logarithmic nature of comprehension. Or, barring this, the individual must be allowed to pursue the study of mathematics on an individual basis--thus making his own allowances for storage time.

Individual Studies

Shortly after the conclusion of the experiment, and in the course of individual interviewing in connection with

another test, the subjects were questioned concerning the mental representatives which they employed for topological concepts and the difficulties which they had encountered in comprehension. Two specific questions were put to each subject:

(a) When you hear the word 'set' (subtest, (V)-space, limit point), what is the first impression (or picture) that comes to mind?

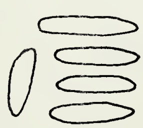
(b) What happens to this 'image' (picture) when you continue to direct attention (concentrate) to it?

The answers which were given most frequently for the first three concepts are listed in the following Table 27. The subjects were also asked to provide sketches for first impression and continued attention 'images'.

The main features of the psychological comprehension of topology are exhibited in this analysis. We notice particularly, the first general impressions giving way to specific representatives, the general awareness of movement and the development of particular representatives (as for example, the 3-element topological space). Which representatives are stored, of course, depends on the examples used by the demonstrator. It is likely that if a two-element topological space had been employed consistently in demonstrations, it would have been stored as the representative for topological spaces.

Table 27

Representatives for Elementary Topological Concepts

Concept	First Impression	Continued Attention	Mathematical Definition
Set	A vague impression of a cluster of points. 	Gradually the identity of a single point emerges--usually visualized as a small circle. 	Undefined.
Subset	A small cluster embedded in or surrounded by a smaller cluster. 	A sensation of movement, picking out, (sometimes pointing to) specific elements in a set. 	Undefined
(V) Space	An impression of a horizontal and vertical configuration of points. 	1. A horizontal listing (movement) of a few elements (usually three) (a, b, c). 2. The vertical listing of each element (a, b, c) a b c 3. The listing of elements of the horizontal set after each element in the vertical set. (a, b, c) a = (a,b), (a,c) b = ( ), ( ) c = ( ), ( )	A set K of elements in which with each element a there is associated a certain class of subsets of K called neighbourhoods of a.

A further investigation was made concerning the representatives utilized by the subjects in Group A who had succumbed to the logical presentation. It occurred almost without exception that the point of breakdown in comprehension coincided exactly with the stage at which adequate representatives failed to be formed--although a correct verbal enunciation of a definition or result sometimes continued beyond the comprehension breakdown point.

This brings into focus the often discussed connection between verbalization and mathematical thinking. In this admittedly limited and special case, the verbal definition performs a 'steering function' with respect to the activation and manipulation of representatives, so that while the verbal definition may be carried along in memory, it is the representative which enters into extensions or elaborations of the definition or results. Without adequate representation, the verbal statement is entirely inadequate.

Summary

A preliminary distinction was made between 'lecturing' and 'teaching'. The former is concerned with a step-by-step logical development of a deductive sequence; the latter implies a correspondence between the demonstration rate of the teacher and the comprehension rate of the student. There is undoubtedly a continuum here in the degree

to which the comprehension problem of the subject is ignored.

For purposes of experimentation, the effect of demonstration at two points on this continuum were examined in connection with the introduction of some elementary concepts of general topology. The first of two groups of grade 12 students (who had been matched in intelligence and general mathematical accomplishment), were presented with the material in a manner which would approximate to a 'superior' mathematics university lecture. As might have been expected, a logical development which was essentially linear in time led to a rapid decline in comprehension over the whole group. The chief difficulty seemed to be that inadequate time had been allowed for the storage of representatives--although verbal definitions were often in evidence beyond the last point of comprehension.

It was then decided to see what kind of time allotment was necessary to maintain a fixed level of comprehension in the group (75%). The development proceeded as before, except that adequate allowances were made for storage, and great care was taken to supply and reinforce representatives. In this case, the comprehension time curve seemed to be roughly exponential over the first few concepts.

The experiment, crude and unstatistical as it

undoubtedly is in some respects, nevertheless provides many points of agreement with the theory. It shows the importance of representatives and adequate storage time, and it suggests that the exponential time function of the 'operation-storage-operation-on-storage' sequence (which was derived in rather artificial situations) may be valid in more complex mathematical systems. In any case, it indicates that the assumption of linearity between demonstration and comprehension times--an assumption on which the mathematics lecture is based--is psychologically invalid.

CHAPTER XI

PROBLEM SOLVING UNDER FINITE UNCERTAINTY
THE GEOMETRY STUDY--PART I

The geometry study investigated problem-solving behaviour over a period of 11 weeks in order to test the hypothesis, advanced on the strength of the energy theory, that some kind of upper bound would be exhibited. In other words, it was postulated that when an individual attempts a series of problems of increasing complexity, the division of available attention-cathexis among an increasing number of alternatives would eventually render the individual incapable of dealing with this complexity. In this case, it would seem likely that he would either 'give up' or behave 'illogically' by ignoring some of the alternatives, thereby decreasing the subjective problem difficulty.

Moreover, if the upper bound is early established, it would allow prediction of subsequent performance in problem solving.

At this point, it may be wise to examine the notion of prediction a bit further. The first point which must be made is that there are several kinds of prediction. The most common type uses the multiple-correlation technique to predict performance in a given area from a set of characteristics (scores) measured prior to the behaviour being

predicted. A second variety, uses predictors measured at the same time as the behaviour under question. A third method uses a sampling technique--and predicts later behaviour from initial behaviour of the same kind.

Prediction of the first kind is undoubtedly the most valuable--provided that valid measures of the criterion are available--for it would allow a before-hand selection of students who are likely to succeed at, and who therefore should attempt the various school subjects. However, the author, arguing from evidence to be presented later, doubts the existence of high a priori correlates of geometry problem-solving ability. Again, while it seems possible to find 'contemporary' correlates of an ability which is measured after considerable experience in an area, these correlates are not necessarily good predictors, because they have been measured at the wrong end (with respect to time) of the behaviour in question.

The 'sampling' kind of prediction would allow the teacher to make adjustments (such as ability grouping) within the class, after assessing ability through its early manifestations. The practicability of such predictions depends on the existence of early-manifested and stable upper bounds.

We postulate, then, that in a long-term situation, the individual tends to find that his performance relative to

the class attains a steady value. Thus guessing (or 'face-saving') patterns would tend to be relatively stable over long periods of time. Moreover, the initial upsurges in motivation which often occur at the beginning of the term, soon give way to the long-range motivation which the student constantly brings to the academic side of his life. In summary, then, there was reason to believe that the total interplay of forces and pressures would cause behaviour to stabilize to the extent that it would become predictable.

1. Objectives of the Geometry Study. The main objectives of the geometry study can now be listed as an attempt to answer the following questions:

(a) Does the student reach a stable level of performance, so that it is possible to say that student X has a chances in b of solving problem Y in a given classroom situation?

(b) Is it possible to construct a scale of difficulty for geometry problems against which problem difficulty levels can be measured?

(c) At what stage is such prediction possible?

2. Procedure. The geometry study was comprised of two phases.

(a) A pilot study employed four students for a period of time from 20-25 hours each. This semi-clinical investigation was conducted with a view to determining the course of mental processes followed in the actual solution of geometry problems.

(b) The autumn investigation encompassed an 11 week period, beginning with the opening of the school in September, 1958 through to mid-November. Details of the autumn investigation are elaborated below.

3. <u>Population</u>. The subject population consisted of two grade 10 groups (N=69) and one grade 12 group (N=37), yielding a total population of 106 subjects. The grade 12 group[1] had received instruction in grades 9 and 10 from the author, and so were well known to him. Data concerning age, IQ, and scores on standardized tests were listed in an earlier chapter. (p. 98).

Euclidean plane geometry is studied for something more than the first half of the grade 10 course and again constitutes the course of study for the entire grade 12 year. The geometry programme is illustrated in Table 28.

The experimental geometry course devolves largely around the use of the ruler and compasses, the acquisition of a geometric vocabulary and the discovery of some elementary geometric facts through measurement.

The author took the regular classes in geometry with

---

[1]In Ontario, the secondary school curriculum is of 5 years' duration; thus the grade 12 students are in their second-to-last year, and are not subject to external departmental examinations.

the grade 10 and grade 12 groups. Each class received one 35 minute lesson a day, five days a week. In all, the learning period extended over 50 lessons, or approximately 30 hours. During this time, the author acted with the full authority of the regular teacher with respect to assignment of marks, disciplinary measures, and general classroom management.

Table 28

Geometry in the Ontario High School

Grade	Duration	Subject Matter
9	3 weeks	Experimental Geometry
10	4-5 months	Euclidean Geometry--Books I and II
12	9 months	Euclidean Geometry Review Books I and II, Books III, IV, and V.
13	9 months	Analytic Plane Geometry.

4. Problems Used. The author felt obligated in some respect to follow as closely as possible the normal course of study for both grades. For, not only would a disruption of 50 lessons duration be intolerable to school authorities, but the relevance of research findings to the ordinary classroom teacher depended upon close correspondence to their conditions of teaching, especially with regard to subject matter. Thus, the problems were selected, for the

most part, from the textbook in current use in Ontario (Petrie, 1954). In the pilot study, all the problems following each proposition were employed and analysed for suitability in the main study. Some of these were later deleted, largely because they were duplications of earlier ones or involved unusual constructions. The final battery numbered 86 problems for the grade 10 group and 92 problems for the grade 12 group. Moreover, to facilitate comparison between the grades, over half of the grade 10 problems appeared in the grade 12 battery.

5. The Course Outline. Some of the propositions of Book I contribute in no substantial way to problem solving. In particular, the 'construction' propositions are of this kind. These were dealt with rather lightly and no problems were assigned after them. The complete system of axioms, definitions, and propositions is given in Appendix D.

The system does not--as is well known--give a complete mathematical axiomatization of Euclidean geometry. Nevertheless, it is the system used in the schools, and for that reason was employed in this study.

6. Teaching Procedure. The teaching method followed the normal classroom approach. After a preliminary discussion of the axioms and the deductive method in general, each proposition was then taken in turn and its proof

developed with the assistance of the class. Considerable care was taken to explain and reinforce the major reference points (i.e., the propositions) and tests were administered to ensure their comprehension. This was considered important because the theoretical discussion of problem solving assumed 'comprehension' of reference points.

After each major proposition, a battery of problems was administered, beginning with very easy and immediate applications of the proposition, and advancing in difficulty until a point was reached where no solutions were forthcoming from any member of the class.

The problems were presented on mimeographed sheets of paper on which a diagram, a statement of the given facts and a statement of the required facts were given. The student was also required to record the beginning and finishing time for each problem in locations provided for that purpose. A typical sheet is shown in Appendix D.

The problem battery usually extended over more than one period. The student worked at his own pace, and it was frequently pointed out to him that the number of problems attempted was not so important as the correctness of the solutions which were offered. Thus the student was in no ostensible way competing against time limits, nor was he 'pressured' to keep up to other members of the group. In

fact, every attempt was made to minimise guessing behaviour by concealing the differences in rate of performance by different subjects in the group.

The work of each student was examined and marked daily so that evidence of upper bounds was readily available. The sequence of problems was continued until it was evident that each student had reached a point where he was solving problems incorrectly or had given up. At this point the sequence was terminated and a new proposition introduced.

Each period during the problem session was introduced by considering problems which had been attempted by all the students on the previous day. All problems were solved using the 'strategy' previously defined, as illustrations of its method and certainty of success.

To further ensure the availability of the problem-defining strategy, each student was provided with a 'problem-solving sheet' (Appendix D). Here the steps of the strategy were explicitly stated and a table showing the permissible propositions for each type of decision was given. The subjects were allowed, and expected, to keep this sheet available and to make use of it during the problem-solving sessions.

The author had observed in his own teaching, that the cumbersome notation employed--this is particularly true of angles--caused an unnecessary extra source of confusion.

Moreover, the traditional demonstration of the 'proof' which requires the elaboration of each item compared, soon involves so much sheer writing that the student may spend by far the greater part of the allotted time writing down proofs rather than thinking. For the purposes of the experiment, an abridged notation was used which substituted algebraic symbols whenever possible and which required only the elaboration of the logical steps in the proof.

Various means were employed to maintain optimum motivation including the usual appeals to term marks. In addition, the problem series began with problems of difficulty index zero, so that all students had at least periodic experiences of success. The author also wrote remarks on the solutions which were intended to convey the impression that the student's work was satisfactory and that continued effort would be rewarded with at least a pass mark, irrespective of actual results. Fortunately, the author was able to make good this promise.

No outside assignments were given, nor were textbooks of any kind employed in the class, or available to the students. Thus, the problem-solving activity of the students was confined to the classroom. At all times, the students were impressed with the necessity of working by themselves and asking for help was systematically discouraged.

7. Preliminary Scoring and Tabulation. The daily work of each student (each problem) was analysed and entered on a large sheet under the following headings:

(1) Solution Correct.
(2) Solution nearly correct: one minor error.
(3) Solution partly correct: one major error.
(4) Solution incomplete, but correct as far as given.
(5) Solution incomplete: partial solution contains errors.
(6) Incorrect solution: complete and contains two or more major errors.
(7) No solution offered.
(8) Time required for any of the above solutions.
(9) Original or unusual solutions.
(10) Logical Errors:
  (a) Assumption of facts not given.
  (b) Misuse of proposition: stating proposition where required facts are not given.
  (c) Use of an illegitimate authority; generally a proposition not yet taken.
  (d) Assuming what is to be proved.
  (e) Illogical conclusion from two given premises.
  (f) Bluff: an argument which concludes with the required facts but which is not based on a logical analysis.
  (g) Omission of necessary steps.

Thus, for each individual, and each class, a complete day-by-day record of problem-solving performance was available. In all, approximately 9,000 problems, representing investments of the students' times ranging from a few seconds to half an hour, were analysed.

Measurement of Problem Difficulty

Two measures of problem difficulty have already been indicated and demonstrated. It will be the purpose of this section to examine these further in the light of the present

investigation.

The following table shows the problem-solving performance of a grade 10 student over the four consecutive quarters of the experimental period. The entries represent the percentage of problems in each cell which were correctly solved by the subject in question.[2]

Table 29

Stability of Problem-Solving Performance

	Index of Problem Difficulty			
0	1	2	3	4
First Quarter	100	67	25	0
Second Quarter	100	75	33	0
Third Quarter	100	75	25	0
Fourth Quarter	100	75	33	0

It is clear that this subject exhibited a well-defined upper bound since he solved no problems beyond difficulty index 3. However, it is important to note that not all problems <u>below</u> the upper bound were solved; in fact,

---

[2]This profile has been chosen for illustration purposes because it is exceptionally well-defined. A measure of the stability of the scores at each level of difficulty is given by the intra-class coefficient of correlation (Haggard, 1958) which has a value of .97 in this case.

the percentage tended to increase gradually over successive categories of decreasing problem difficulty. It is also evident that in prediction of performance in this case, something more than a statement of upper bounds is possible, viz., that the student is likely to solve all problems in the interval 0-1, approximately three-quarters of the problems in the interval 1-2, approximately one-third of the problems in the interval 2-3 and no problems above difficulty index 3. It may well be that for educational purposes (in the sense of classroom grouping), this student would be considered to have a 'practical' upper bound of 2, since he solves less than one-half of the problems in subsequent categories.

Another consideration centres around the 'strength' of the measure of problem difficulty. In the profile above, for example, problem difficulty was assessed from the difficulty index (percentage pass) converted to a standard score. If the first-half and last-half upper-bound scores are correlated for all members of the group, the coefficient which results is a measure of the stability of the individual's performance relative to the group. This does not exclude the possibility that trends are present within the group, and so the existence of a high coefficient may be taken as a necessary, but not sufficient condition for the

existence of upper bounds.

While this type of stability may be adequate for dealing with individual differences within the classroom, the existence of upper bounds requires that problem difficulty be measured independently of the group. For the external quantification of problem difficulty, we must resort to model construction, i.e., we suppose that the individual solves the problem in a certain manner, and on this assumption, we calculate the number of operations performed, and hence infer difficulty. We must realize at the outset, however, that models are only approximations to reality. In the attempt to quantify, we must necessarily gloss over many of the subleties of human thinking.

Two models have already been considered: the deduction model, and the uncertainty model. They will be examined in turn in the following pages.

1. The Deductive Measure of Problem Difficulty. Since the deductive measure--i.e., the number of statements required in the proof--is used as a basis for assigning marks by many high school teachers, it is well to estimate its validity as a measure of problem difficulty. For this purpose, the deductive lengths of the 86 problems of the grade 10 battery were determined, and profiles constructed similar to the one shown above. The determination of deductive length was made on the basis of the listed axioms and

deductions of the system (Appendix D) and so represents a refinement of the teachers' measure. The correlation between deductive length and the standardized, experienced difficulty index amounted to .66 (N=86). Moreover, the consistency of the practical upper bounds (defined in this case as the level in the difficulty scale above which the individual will solve less than one problem in two) is indicated by the first-half, second-half coefficient of .84.

While neither correlation is spectacularly high, there is some indication here that the traditional method is moderately valid, and there is a strong indication that upper bounds--even though defined in this somewhat crude measure, do exist.

2. The Uncertainty Measure of Problem Difficulty. In the development of the uncertainty model, it was seen that some assumption had to be made concerning the subjective probabilities. Each assumption gives rise to a different variation of the uncertainty model.

(a) Model A: $w_{ij}=1/N_{ij}$

In this case, the subject gives equal weight to the various propositions. It has been shown that this tends to maximize the difficulty of each decision.

(b) Model B: $w_{ij}=p_{ij}$

In this case the subjective probabilities are equal to the objective probabilities for the set in question. This would seem at first glance to be an unrealistic assumption, because it can be argued that the subject beginning a set of problems cannot know what the relative frequencies will be until he has, in fact, finished the set.

Offsetting this, however, is the effect of 'proposition dominance'. The fact that a particular proposition precedes a set of problems tends to give strong emphasis to the particular proposition as the proper choice for a certain type of alternative. Thus, a strong subjective probability for the particular proposition is introduced.

It also happens that the set of problems following the proposition does tend to use this proposition extensively, so that p_{ij} is high for the proposition in question. Thus there occurs a fairly close matching[3] of p_{ij} and w_{ij}.

<u>Model C</u>: w_{ij}=relative frequency up to the problem in question. This model appears quite plausible on the surface. It will be shown however, that it leads to some

---

[3]There seems to be, in fact, a definite psychology of problem construction. The test constructor aims primarily at providing a sequence of increasingly difficulty applications of the proposition in question (thus 'proposition dominance'), and beyond that he invokes previous propositions in a more-or-less random fashion. The matching of the subjective and objective probabilities occurs because the student acquires a sensitivity to this method.

unrealistic results in not allowing for 'proposition dominance'.

The models were compared in the pilot study by considering the table of reference of a student who is just about to attempt the set of problems following Proposition III.

Following Model B, the relative frequencies for the set of problems is given in Table 30.

Table 30

Subjective Probabilities for Uncertainty Analysis
Model B

To Prove	Axiom A	I	II	III	
$\angle = \angle$	.14	0	.10	.76	Theoretical difficulty of angle decision =1.03 units[a]
$\ell = \ell$	0	0	1.0	0	Theoretical difficulty of line decision =0.00 units
$\triangle \equiv \triangle$	0	0	1.0	0	

[a]The difficulty of the decision is calculated in each case by means of the uncertainty formula

$$U=-\sum p_i \log_2 wi$$

This illustrates very clearly the phenomenon of 'proposition dominance'; here, Proposition III dominates

the angle decisions in the set of problems following it.

Suppose next that the individual is operating according to Model A. He would then possess the following table of reference.

Table 31

Subjective Probabilities for Uncertainty Analysis
Model A

To Prove	Axiom A	I	II	III	
$\angle = \angle$	.25	.25	.25	.25	Theoretical difficulty of angle decision =2.00 units.
$\ell = \ell$	0	0	1.0	0	Difficulty of line decision =0.00 units
$\triangle \equiv \triangle$	0	0	1.0	0	

We notice here, the expected increase in the difficulty of angle decision.

Again, suppose that the individual is operating according to Model C, in which case he possesses Table 32 at the beginning of the exercise.

The relative frequencies under III are all zero because Proposition III has not yet been used in a choice decision. In this case, the quantity

$$U = -\sum p_i \log w_i$$

$$= -(.14 \log .53 + 0 \log .314 + .10 \log .151 + .76 \log 0)$$

$$= \infty$$

Table 32

Subjective Probabilities for Uncertainty Analysis
Model C

To Prove	Axiom A	I	II	III
$\angle = \angle$	.515	.304	.152	0
$\ell = \ell$	0	0	1.0	0
$\triangle \cong \triangle$	0	0	1.0	0

Thus, if we proceeded with this set of probabilities, the first problem would have a calculated difficulty value of infinity. If the entries in the probability table were assumed to be the relative frequencies at the end of the first problem, then a similar calculation gives:

$$U = 4.20$$

which is again, far too high. If we assume that probabilities are equal to the relative frequencies at the end of the set of problems, then,

$$U = 1.80$$

which lies closer to the Model B value of 2.00 than the Model C value of 1.03. This model completely ignores the

effect of proposition dominance which proved to be a very real effect in the problems under consideration. This model was therefore dismissed as inadequate to explain the observed difficulty.

Models A and B were compared in terms of the accuracy of their assessment of problem difficulty as compared to experienced difficulty. In both the summer and autumn studies, Model B proved to yield results corresponding more closely to experienced difficulty and it will be used in all subsequent discussions.

An estimate of the validity of the uncertainty model is given by a correlation of .87 with experienced difficulty at the grade 10 level[4] (N=86). A first-half, second-half profile correlation of .90 indicates that the practical upper bounds (defined as for the deduction model) exhibited considerable stability, and it is evident that this measure offers some improvement over the deductive measure. The circumstances which favour one measure over the other are considered in the following section.

The Pilot Study: Analysis of Individual Problems

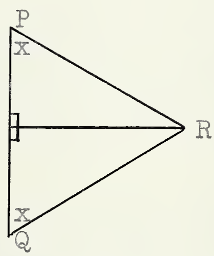
The pilot investigation was conducted before the data allowing estimates of model validity were available. In

---

[4]The deductive and uncertainty measures are substantially related, yielding a product-moment r of .70 for the grade 10 battery (N=86).

many ways, the study of individual solutions corroborated this later data and showed the particular areas where the two models failed to correspond to the method actually employed in solution. Some particular instances are considered below.

1. Solutions to the easier problems, especially those which followed immediately after a proposition, tended to deviate from the analysis strategy. The following simple example followed immediately after Proposition XVIII.



Given: $\angle P = \angle Q$

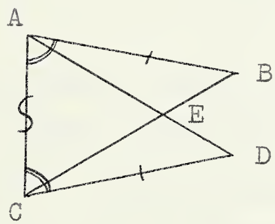
Prove: $PR = QR$

Fig. 24. Geometry problem, illustrating a deviation from the analysis strategy.

Here, the analysis strategy would require the individual (under Q_2) to consider all the possible propositions which may be used to prove that one line equals another, and to make a choice among the three resulting alternatives. However, the fact that this problem followed directly after Proposition XVIII usually led to an immediate enunciation of this authority as proof. There was little evidence of a consideration of alternatives in the verbalized

solutions of problems of this kind. Generalizing somewhat, the analysis approach was seldom in evidence until a gap length of two or three units was involved.[5]

2. Another deviation from the analysis strategy was seen in the tendency toward 'spontaneous' deduction--without regard to its relevance to the solution--which often took place when a diagram occurred which obviously belonged to a particular proposition. The example below illustrates this point.[6]



Given: Diagram as marked.

Prove: $\angle BAE = \angle DCE$

Fig. 25. Geometry problem, illustrating deviation from the analysis strategy.

This problem occurred after Proposition II; its solution requires the congruence of triangles ACB and ACD.

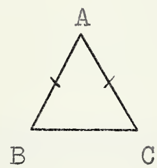
---

[5]The uncertainty measure lends to compensate to some extent for the increased number of prepositions (T_i) through the principle of 'proposition dominance'.

[6]Again, this situation is to be thought of as a deviation from the analysis strategy which is nevertheless adequately dealt with by the uncertainty formula. It happens that spontaneous deduction usually occurs in conjunction with a fairly complex diagram, and in this case the uncertainty measure tends to be high as well.

A surprisingly large proportion of students immediately marked ∠AEB=∠CED, although this contributes nothing to the solution. In fact, it further distracts and divides attention from the proper analysis.

An even stronger urge to make the 'obvious deduction' seems to attend the configuration,



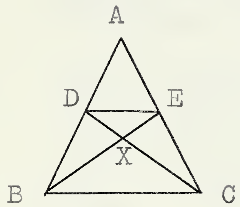
in which the angles B and C are almost invariably and immediately marked equal (in complex diagrams) whether or not this contributes to the solution. It is not suggested here that the practice of 'spontaneous' deduction is always deleterious; it often happens that if such results are marked on the diagram, without obscuring its essentials, the net result is a decrease in the gap length.

Having considered two cases which show the analysis strategy to disadvantage, we next consider a case where the superiority of this method over the deduction model becomes obvious.

3. The uncertainty estimate of difficulty is invariably closer to experienced difficulty, if there is considerable complexity of geometric configuration in the problem

under study.

This is particularly true in cases such as the following, in which triangles overlap. The problem occurs after Proposition II.



Given: AB=AC

AD=AE

Prove: DC=BE

Fig. 26. Geometry problem, illustrating the superiority of the uncertainty measure.

The deductive proof inself involves only two statements:

Proposition II $\longrightarrow \triangle ABE \equiv \triangle ADC$

Definition of Congruence $\longrightarrow$ BE=DC

However, the uncertainty analysis shows clearly the effect of the multiplicity of triangles, for any of 5 pairs of triangles may be considered as a possible choice to prove the required fact. A recorded verbal solution will be valuable at this point. The solution is set out as it appeared on the sheet employed by the author to record solutions in the pilot study.

The subject here was a fifteen-year old boy, with a Wechsler IQ of 112--a boy who proved to be a slow, but effective problem solver.

Time	Subject's Behaviour	Comment
0-20 sec.	Reads problem aloud. Begins by tracing out lines to be proved equal. Looks at problem-solving sheet. "I need congruent triangles to use Proposition II."	Following analysis strategy.
30 sec.	Traces lightly over $\triangle DXB$ and $\triangle EXC$. Marks $\angle DXB = \angle EXC$. Pauses, looking for sides. "I'd need DX or BX,--(Pause), but they're part of what I'm supposed to prove. Let's see, what else?"	Considers the most obvious alternative. Dismisses it temporarily while going on to other triangles. Spontaneous deduction exhibited.
	Traces out $\triangle BDE$ and $\triangle DCE$, $\triangle DBC$ and $\triangle EBC$, $\triangle DXE$ and $\triangle BXC$. (Pause).	Enumerating alternatives.
90 sec.	Returns to $\triangle DBC$ and $\triangle EBC$. "Let's see--I've got BC common and BD (pause) equals EC." (Demonstrates and marks on diagram.) "Maybe if I can prove $\angle B = \angle C$ (pause)--they both have two smaller pieces."	Examines another alternative and makes some deductions which do not contribute to the solution.
150 sec.	After a pause, considers $\triangle BDE$ and $\triangle DCE$. Traces out known sides. "We still don't know the angles."	
	Traces out $\triangle ABE$ and $\triangle EBC$. Turns page sideways. "No." Very long pause.	
240 sec.	"Well, let's see--I want to prove a line equals a line so I need Proposition II, and congruent triangles. There's	Comes back to analysis strategy.

Time	Subject's Behaviour	Comment
	$\triangle$DBC and $\triangle$EBC, etc., (lists all he had before and pauses after each, reconsidering its possibilities). "There must be some more." (Draws each pair separately.)	
480 sec.	Locates $\triangle$ABE and $\triangle$ADC and solves at once.	Solution obtained in 8 minutes.

The impressive thing about this solution was the return to the analysis strategy when an equilibrium point was reached, and the inability to completely dismiss the various alternatives from mind. The uncertainty concept seems to give an accurate picture of the difficulties encountered in this case.

Prediction Diagrams

Since both measures of problem difficulty have some validity and tend to complement each other to some extent, it would seem conceivable that some combination of them could be used to advantage, particularly with reference to prediction. One-dimensional prediction graphs are common enough in psychometric measurement (Wesman, 1949); they consist essentially of a scattergram of scores, a prediction variable plotted against the criterion variable. This idea

was extended in the present case to two dimensions--the 'uncertainty' and 'deductive' measures of problem difficulty were used as the rectangular co-ordinates of each problem for the subjects of the pilot study (Fig. 27). The correct solutions are recorded in blue, the incorrect in red, and the time of solution is written under each problem. The numbers recorded along the axes represent solution times for correct, incorrect, and combined questions in the given row or column respectively. The percentage of correct solutions in each row and column is also given.

Upper bounds in both the uncertainty or deductive measures are in evidence here; in addition, the profile shows a regular decrease in the proportion of problems solved as the difficulty level rises. Judging from the solutions offered, the area of 'mastery'--i.e., the region of the plane in which the student can solve a majority of problems--seems to be roughly within the region defined by the equation,

(Deductive measure) + (Uncertainty measure) =Constant

This would indicate that a suitable estimate of problem difficulty would be given by the sum of the two separate measures. Such a measure would tend to rectify the tendency of the uncertainty measure to overestimate the difficulty of problems of small gap length, and the tendency

			0–1	1–2	2–3	3–4	4–5	5–6	6–7
		8				IV 180	IV 480		S 840
		7					IV 210	S 600	
0	430 430	6				IV 80	II 540, IV 510		
28	180 300 266	5		II 150	I 90, O 240	III 120, IV 360, IV 300	II 600		
55	178 253 209	4		II 170, II 420, IV 30	I 5, I 35, III 480, III 90, III 105, IV 180	II 200, II 150, II 150, III 240, III 180, III 540, III 420	IV 140, IV 480		
83	87 370 134	3	O 20, O 30	I 15, I 10, II 10	III 65, III 330, III 240, IV 35	II 410, II 270, IV 180			
100	60 60	2	O 20, O 15, II 10, II 30, II 65	I 10, I 15, II 70, II 165, III 45	II 110, III 120	IV 140			
		1							
			0	1	2	3	4	5	6 7
% Correct			27 27 100	74 160 87 85	149 159 152 71	175 297 240 47	140 595 544 11		

Fig. 27. Geometry solving performance over the first four propositions.

of the deductive measure to underestimate the difficulty of problems involving complex figures.

The study of the time required for solution also offers some points of interest. The average solution times for successive intervals in both dimensions show regular increases; moreover, the increases for the correct solutions seem to be approximately linear on the uncertainty scale. (This becomes more evident when several profiles are put together.) If true, this would indicate that the raw uncertainty measure offers an interval scale of difficulty (if we assumed that problem difficulty is proportional to time required for solution).

The diagram can be used to predict the probability of success on future problems. If the measures are weighted equally, the expected value of the probability of success for each cell may be obtained by averaging the percentage passes in the row and column corresponding to the given cell. Similar calculations can be made for the expected time.

The success of such predictions can be assessed from Fig. 28, where the solutions for 28 problems following Proposition IV have been recorded. The outcomes of the four easiest and the 8 hardest problems were correctly predicted. Moreover, in the 'region of doubt'--i.e., near the upper

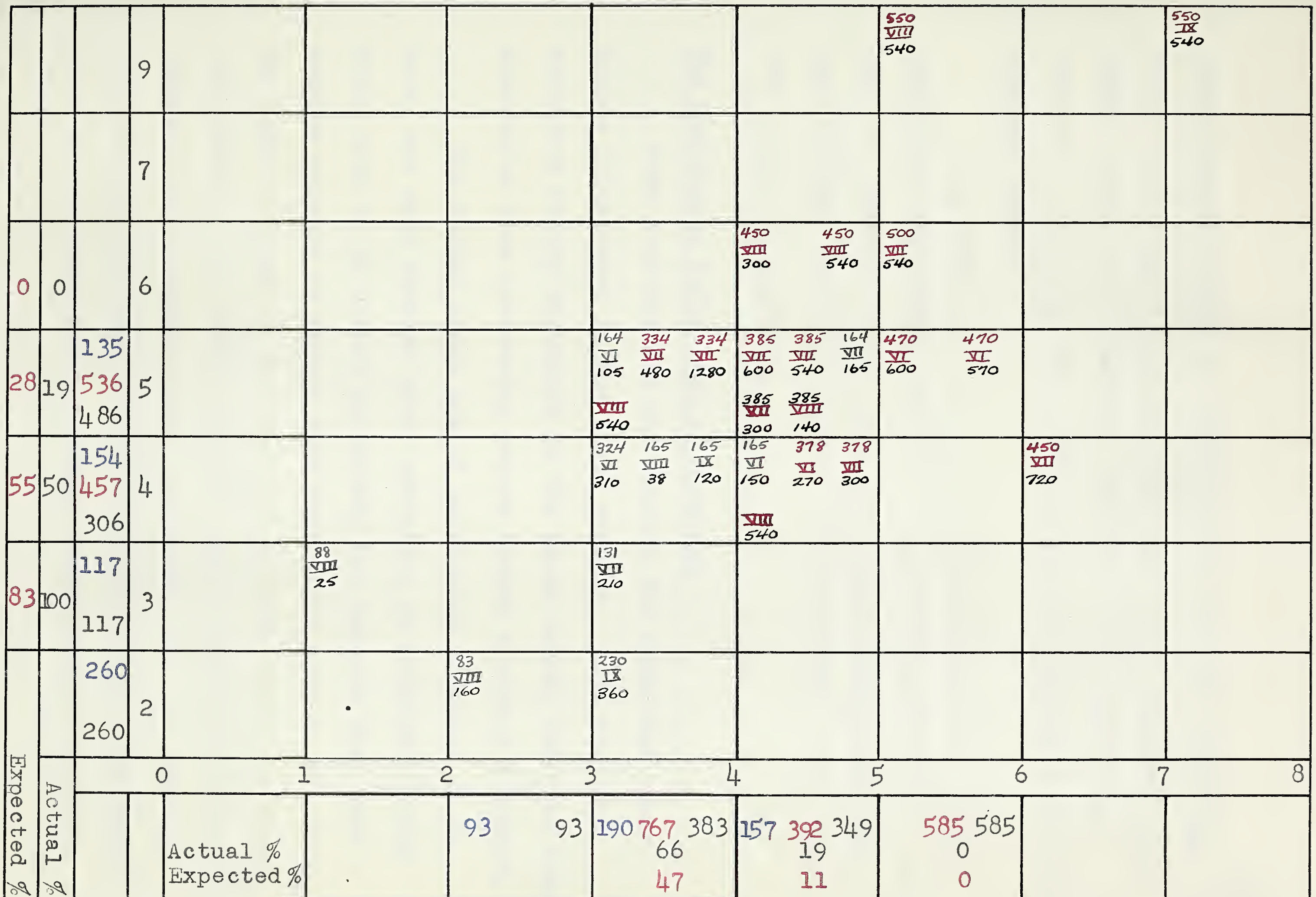


Fig. 28. Use of combined uncertainty and deductive measure in the prediction of geometry performance.

bound (square region in centre)--the predicted percentage of passes (40%) agrees closely with the actual observed value (37%). Again, the predicted times for both correct and incorrect solutions show a fair amount of agreement with the observed values.

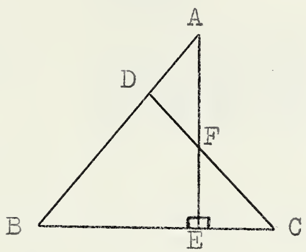
In general, it is possible to say that this student exhibited predictable stability in geometry performance, and while the prediction was made after four propositions had been covered, prediction after two propositions would have been nearly as accurate.

The Continuous Refinement of Strategy

Some profiles did not exhibit the clear-cut upper bounds illustrated in the above example. Some students exhibited steady increases in the upper bound, measured over successive time intervals; others showed a steady decrease.

The former group are of particular interest to us here, not only because their potential as problem-solvers would seem to be almost unlimited, but because they have somehow managed to escape from what might seem to be--from the point of view of the theory--necessary limits to their performance. In a later chapter, some personality characteristics of this group will be considered; for the present, the recorded solutions of another problem will be examined, since it reveals how the superior problem-solver modifies

the static strategy to his own advantage. The solution will be set out, as before, in the fashion in which it was recorded.



Given: BD=DC

$\angle AEB = 90^{\circ}$

Prove: $\angle DAF = \angle DFA$

Fig. 29. Geometry problem 4.

Time	Subject's Behaviour	Comment
0-10 sec.	Reads problem. Marks DB and DC equal, also marks $\angle B$ equal to $\angle C$ ($= x$).	The spontaneous deduction exhibited here, turns out to be useful in this case.
20 sec.	"Prove one angle equal to another--it would not be congruence here--or parallel lines--it must be angles within a triangle. Proposition XVI or III." (Pause),--"XVI in this case."	Seems to have divided angle decision into three general categories.
40 sec.	Marks $\angle DAF=a$, $\angle DFA=b$, then marks $\angle EFC=b$.	'Directed' deduction.
50 sec.	"To compare b and a, we have to get them in another triangle." Traces out triangles with pencil.	

Time	Subject's Behaviour	Comment
	"$a+x=90^\circ$ in $\triangle ABE$."	
	Pause.	
70 sec.	Marks $\angle FEC=90^\circ$. "In $\triangle FCE$, $b+x$ equals 90°. Therefore $b+x$ equals $a+x$ and a equals b."	More 'directed' deduction. Solution in 85 sec.

The subject in this case was an extremely bright 15 year old boy, and the problem occurs in the exercise following Proposition XVIII. It appears that the subject was still following the main path of the analysis strategy, but that he had effected further division within the general categories of the basic strategy provided him. Thus, instead of a general 'angle decision', he employed three angle decisions based on congruency, parallelism, and relations among the angles of a triangle. Moreover, he seems to consider only the latter possibility--we might say that he had high subjective probability for it--so that, in effect, the choice narrowed from 7 to two propositions. Thus, one characteristic of the problem solver who makes continued progress in geometry, is that he counteracts the increasing multiplicity of propositions by organizing them into categories which he associates with certain geometrical

configurations.

This principle of increasing complexity of organization can be set against claims for any genuine discovery. Very few instances of originality actually occurred in the solution of geometry problems in the study. In fact, the occurrence of the necessity for any new 'type' of proof--such as equality by congruence, or indirect proof--invariably rendered the pupil incapable of obtaining a solution. What we do find, in so-called 'original' proofs, is usually a transfer of technique.

Another important aspect of this proof is the occurrence of 'directed' deduction. In this problem, marking $\angle B$ equal to $\angle C$, $\angle DFA$ equal to $\angle EFC$ and $\angle CEF$ equal to 90°, all contributed to a reduction of gap length and consequently, to a reduction in the number of analysis steps required.

The subject was later asked to explain what prompted him to mark these angles. It happened that the first deduction (viz., $\angle B = \angle C$) was truly spontaneous, accomplished without any particular thought of its consequences, indeed, accomplished before the problem had really been considered at all. This was not true with respect to the other deductions, however. For example, when the subject marked $\angle DFA = \angle CFE$, he had decided that angles DAF and AFD would have to be related to angles in other triangles and he felt that the equality of $\angle AFD$ and $\angle CFE$ had something to do with this,

although he was not certain just what this connection was. Thus, the deduction in question was not undirected (as for example, $\angle DFE = \angle AFC$ would have been) but was directed toward some phase of the analysis.

Two results stand out in this discussion. When we construct models, we are employing a static method of looking at the subject's behaviour. The model does not adequately allow for the incorporation of broad sections of experience into patterns. However, the basic hypothesis concerning the constancy of hypercathexis remains, but the individual who establishes successively higher equilibrium points manages to escape from the constant equilibrium level which would result from a fixed strategy.

While the analysis and deductive methods of determining problem difficulty are to this extent inadequate, since they tend to overestimate problem difficulty for students adopting generalized strategies, yet it will be shown in the following chapter that they are useful in that they do assess and predict problem difficulty for a majority of subjects.

CHAPTER XII

THE GEOMETRY STUDY--PART II

This chapter deals in part with the evidence obtained from the autumn study, which can be construed as supporting the hypothesis that stable upper bounds will occur in the students' geometry solving performance. A preliminary overview will be provided in order to give some measure of coherence, since the chapter might otherwise appear somewhat fragmentary.

We shall first consider the question of the stability of problem-solving performance relative to the group ('weak' stability). Then, the problem of 'strong' stability (i.e., measured independently of the group) will be approached in three ways: 1. 'category' boundedness, 2. the effect of repeating problems, and 3. comparison of the matched grades 10 and 12 groups.

In a final section, the behaviour of one subject on problems which are below, at, and above, the upper bound will be considered.

Consistency of Performance in Geometry

It was mentioned in the previous chapter that a necessary, but not sufficient, condition for the existence of upper bounds would be the consistency of scores relative

to the group. Indices of difficulty (percentage passing) for each problem of the autumn battery were computed for grades 10 and 12, and then converted to σ scores, measured from an arbitrary zero of -2.00σ.

From these scores, profiles were constructed similar to the one discussed previously. It was obvious from an inspection of these proviles that intra-individual performances are highly consistent while inter-individual performances varied considerably. Although intra-class coefficients for the 'best' and 'worst' profiles were significant beyond the .01 level, this does not tell us much, since significance of the intra-class coefficient would be a minimum requirement for relative stability.

It was decided to calculate a product-moment coefficient of correlation in the following way: The term work was divided into four sections, each consisting of 6 levels of difficulty. The percentage of problems solved correctly was computed for each cell. The percentage values were 'smoothed' to offset the effect of arbitrary intervals. The resulting values were treated as the percentage solved at the centre of the interval in question. The interval values were checked until a pair was located in which (a) $X_N > 50$, (b) $X_{N+1} < 50$. This upper bound was obtained by linear interpolation between the midpoints of these two intervals.

Product-moment correlations were calculated between performance over the last half of the term and performance over (a) the first quarter, and (b) the first half. The results for grades 10 and 12 are shown below.

Table 33

Reliability of Prediction of Relative Performance on the Basis of First-quarter and First-half Performance

	Grade 10	Grade 12
First Quarter and Second Half	.80	.89
First Half and Second Half	.92	.93

It is apparent from the table that the performance in geometry is highly consistent (relative to the group) at both grade levels. If we consider the subject's steady performance value to be defined by the work of the second half of the term, then the relative 'steady' position of the individual within the group can be predicted by the end of the first quarter (three weeks) with fair accuracy. Needless to say, prediction at the end of the first half is even more accurate. This means that the teacher who has an accurate record of the s tudents' performance can estimate relative ability within the class with sufficient accuracy

to divide into ability groups on this basis, and can be sure that such groups will be stable over fairly long intervals of time. Two observations may be made concerning the coefficients. In the first place, their substantial size resulted from the fact that there existed an extremely wide variation in the ability to solve problems. At both grade levels, the best student solved approximately 10 times as many problems as the weakest student. (These distributions are reported later.)

The increase in the grade 10 coefficient suggests a preliminary period of adjustment--at least for some students. However, once the adjustment or levelling off is accomplished, the relative performance seems to remain stable over very long intervals of time. Indeed, the product-moment coefficient of correlation between the marks assigned to the grade 12 students by the author two years previously (i.e., when they were in grade 10) with the score earned in grade 12 amounted to .88.

An estimate of the reliability of the whole battery as a measure of relative geometry performance was made by inter-correlating the two halves and correcting by the Spearman-Brown formula. Thus a reliability coefficient of .96 at the grade 10 level (N=69) is obtained. This is not really surprising, since the battery represents from 16-19

hours of the students' time.

Boundedness of Problem-Solving Behaviour

During the autumn investigation, the author had an opportunity to attend a district meeting of high school mathematics teachers and to explain his views concerning upper bounds in geometry performance. Most of the teachers concurred in the opinion that students do exhibit such bounds, but they questioned the practicability, from the teacher's point of view, of expressing upper bounds in "a rather meaningless unit of measurement".

When these teachers discuss the difficulty of a geometry deduction they employ a verbal scale which corresponds in part, to the problem's occurrence in a sequence of problems of increasing complexity and in part to the number of propositions employed in its solution (deductive length). Thus, they speak of 'immediate applications of a proposition' and 'two-step problems' to indicate two orders of increasing problem difficulty.

A combination of the uncertainty and deductive measures was employed to divide problems into four categories to which descriptive terms could be applied which would be meaningful to the teacher. To do this, the two measures for each problem were computed, converted to σ scores, and averaged (the distributions of problems in each measure were

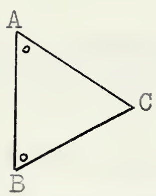
comparable).

A scrutiny of the results revealed two fairly well-defined clusters; category 1, $\sigma \leq -2.0$, which covered the easiest problems, and category 4, $\sigma \geq +2.0$, which covered the hardest problems. The remaining problems were divided into two equal-difficulty intervals: category 2, $-2.0 < \sigma \leq 0$, and category 3, $.0 < \sigma < 2.0$. The four categories can be defined in the teacher's descriptive terms as follows:

1. Category One. The problems in this category could be further subdivided into four grades of difficulty.

(a) Immediate application of a single proposition. Diagram similar to a proposition.

Example: (following Proposition XVIII).



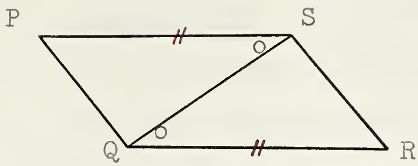
Given: diagram

Prove: AC=BC

Fig. 30. Geometry problem representative of difficulty-category one.

(b) Immediate application of a single proposition, but some part (e.g., a line) necessary for the proposition requires identification.

Example: (following Proposition II).



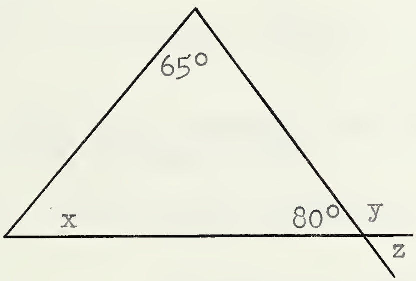
Given: Diagram as marked.

Prove: $\angle P = \angle R$.

Fig. 31. Geometry problem representative of difficulty-category one.

(c) Immediate multiple application of a single proposition.

Example: (following Proposition XVI).



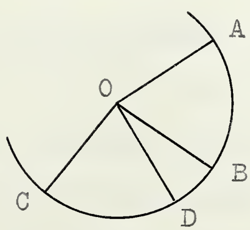
Given: Diagram as marked.

Find: x, y and z.

Fig. 32. Geometry problem representative of difficulty - category one.

(d) A semi-immediate application of a single proposition; some obvious fact needed first.

Example: (following Proposition II).



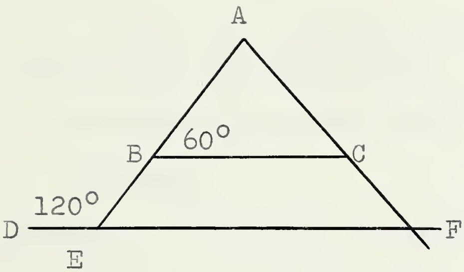
Given: OA, OB, OC, OD, are radii of a circle.
: $\angle AOB = \angle DOC$

Prove: AB=DC

Fig. 33. Geometry problem representative of difficulty - category one.

2. <u>Category Two</u>. The problems in this category were almost all two-step problems (i.e., involving two propositions) in which the application is straightforward and the diagram is not complex.

Example: (following Proposition XVIII).

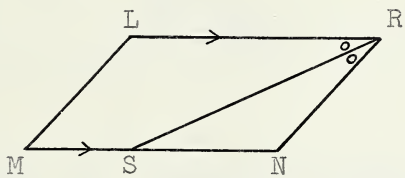


Given: Diagram as marked.

Prove: BC∥EF

Fig. 34. Geometry problem representative of difficulty - category two.

Example: (following Proposition XVIII).



Given: Diagram as marked.

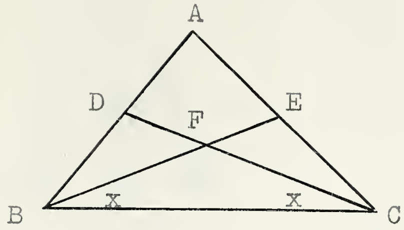
Prove: SN=NR

Fig. 35. Geometry problem representative of difficulty - category two.

3. <u>Category Three</u>. The problems in this category were for the most part, of two kinds.

(a) Two-step problems with some complexity of diagram (especially overlapping triangles).

Example: (following Proposition XVII).



Given: AB=AC

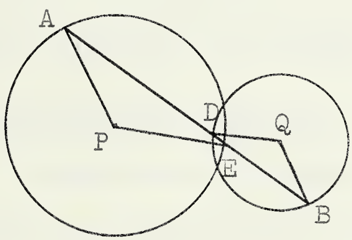
: ∠EBC=∠DCB

Prove: DC=BE

Fig. 36. Geometry problem representative of difficulty - category three.

(b) Three-step problems which are relatively straightforward.

Example: (following Proposition XIV).



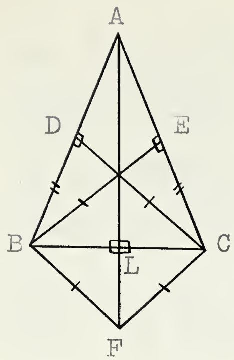
Given: P and Q are the centres of the given circles.
PA, PE, and QD, QB are radii of the circles.
AP ∥ BQ

Prove: PE ∥ DQ

Fig. 37. Geometry problem representative of difficulty-category three.

4. <u>Category Four</u>. Problems were assigned to this category if they were more difficult than those in the previous categories.

Example: (following Proposition IV).



Given: Diagram as marked.

Prove: AD=AE

Fig. 38. Geometry problem representative of difficulty - category four.

It seems apparent enough that the combined measure of problem difficulty did effect a division of problems along a gradient of natural difficulty.

Problem-Solving Profiles

Profiles were again constructed for each student, showing the proportion of problems solved in each difficulty category, and for each quarter of the term work. In order to analyse the profiles in greater detail, each cell was marked '+' or '-' depending on whether the proportion of problems correctly solved in the cell was greater than, or less than .50.[1] The practical upper bound was defined to

[1]The difficulty category containing the upper bound need not have been defined by the constant .50. However, in view of a forthcoming discussion of the illogical aspects of problem-solving demonstrated in the solution of problems beyond the upper bound, it would seem that .50 would be the most natural choice.

lie in the highest difficulty category for which the subject solved correctly more than one-half of the contained problems.

The profiles fell under three general headings which may be described as follows: (1) exhibiting a stable upper bound over the four intervals, (2) exhibiting an increasing upper bound, (3) exhibiting a decreasing upper bound. Certain subtypes which appeared under these general headings are shown in Table 34.

The first row in Table 34 shows the profile types in which improvement over the four quarters is so small that the practical upper bound for the second half-term resides in the same category as the upper bound for the first quarter.

Thus prediction, in this case, would be possible at the end of the first quarter of the work. The other profiles were capable of similar interpretations.

Table 35 shows the complete classification of the 69 grade 10 profiles.

Several observations need to be made with respect to this table.

(a) Predictability of the location of the upper bound at the end of the second half can be made with fairly high accuracy after the first quarter.

Table 36 illustrates the predictability of performance when the upper bound appears in the first category in the

Table 34

Profile Types in the Solution of Geometry Problems

Below 1	1	2	3	4	
1 2 3 4	1 2 3 4	1 2 3 4	1 2 3 4	1 2 3 4	
	Q_1 X - - - H_2 X - - -	Q_1 X X - - H_2 X X - -	Q_1 X X X - H_2 X X X -	Q_1 X X X X H_2 X X X X	Stable upper bound; predictable from the first quarter.
		Q_1 X - - - Q_2 X X - - H_2 X X - -	Q_1 X X - - Q_2 X X X - H_2 X X X -	Q_1 X X X - Q_2 X X X X H_2 X X X X	Increasing upper bound; predictable from the second quarter.
		Q_2 X - - - Q_3 X X - - Q_4 X X - -	Q_2 X X - - Q_3 X X X - Q_4 X X X -	Q_2 X X X - Q_3 X X X X Q_4 X X X X	Increasing upper bound; predictable from the third quarter.
		Q_3 X - - - Q_4 X X - -	Q_3 X X - - Q_4 X X - -	Q_3 X X X - Q_4 X X X X	Increasing upper bound; not predictable.
Q_1 X - - - Q_2 - - - - H_2 - - - -	Q_1 X X - - Q_2 X - - - H_2 X - - -				Decreasing upper bound; predictable from the second quarter.
Q_2 X - - - Q_3 - - - - Q_4 - - - -	Q_2 X X - - Q_3 X - - - Q_4 X - - -				Decreasing upper bound; predictable from the third quarter

Code:
Q_1: Work of first quarter of term.
Q_2: Work of second quarter of term.
Q_3: Work of third quarter of term.
Q_4: Work of fourth quarter of term.
Q_5: Work of last half of term.

Table 35

Classification of Profile Types for 69
Grade 10 Students

	Location of Upper Bound at End of Term					
	Below 1	1	2	3	4	Total
Predictable in first quarter	3	28	9			40
Increasing; predictable in second quarter				4	1	5
Increasing; predictable in third quarter			5	5	1	11
Increasing; not predictable			1	1	3	5
Decreasing; predictable in second quarter	6	1				7
Decreasing; predictable in third quarter	1					1
Total	10	29	15	10	5	N=69
%	14	42	22	14	8	

particular quarter of the term work given by the heading along the top line. The headings along the side indicate the category in which the upper bound will be located at the end of the second half (fourth quarter in the case of column three). Thus, if an individual's practical upper bound was located in the first category at the end of the first quarter, there were 61 chances in 100 that it would still be there at the end of the second half, 18 chances in 100 that the upper bound would be located in category two, and 21 chances in a hundred that the upper bound would be located below category one.

Table 36

Predictability of Location of Practical Upper Bound in Terms of Placement in First Category at End of Successive Quarters

	Located in First Category at End of:		
	1st	2nd	3rd quarter
Stay in 1st	61%	78%	89%
Go on 2nd	18%	22%	6%
Go on 0th	21%	0%	6%

(b) Of the 31 subjects, whose upper bounds fell below category two in the first quarter, only 6 showed significant rise to move up a category. Evidently, poor problem-solving performance is not only demonstrated early, but it remains

consistently poor.

(c) Sixteen of the 21 subjects who showed a rise of one or more categories were above the first category at the end of the first quarter. Similarly, 7 of the 8 subjects who dropped a category were originally in category one. Consequently, if the teacher had assumed that those students whose upper bounds were below category two at the end of the first quarter, would stay there, he would have been correct in 79% of the cases.

(d) There was an increase in the general level of problem-solving over the first three-quarters of the term, since 21 subjects showed a gain of one or more categories, 8 show a loss of one category, and 40 remain constant. Again, many of the 40 who did not show a full category gain, did show a partial gain. For example, the student whose profile is shown in Table 37, seemed to make a partial, slow gain while his practical upper bound remained in category one.

Table 37

Individual Profile Showing No Gain in Practical Upper Bound

	1	2	3	4
First Quarter	.90	.20	0	0
Last Half	1.0	.40	0	0

(e) The overall performance of a large proportion of students was so poor that one can legitimately question the extent to which these individuals made any genuine progress in the ability to solve problems. At the end of the fourth quarter (11 weeks), 33 of the 69 students were solving less than half of the problems in category two, and 10 of these were able to solve less than half of the problems in category one.

In other words, no more than half of the subject population was able to proceed much beyond immediate applications of the propositions. These weaker students did not seem to accomplish much more than to learn the individual propositions and where they applied.[2] Even so, the author would claim that the performance of the experimental group, because of the precautions taken to provide a strategy and maintain contact with the students' progress, was superior to the performance of students in comparable grade 10 groups in Ontario. Unfortunately, the teacher in the ordinary classroom has no opportunity to properly assess the problem-solving activities of his students.

(f) The 10 students who retrogressed in performance present an interesting group. One would be tempted to say

---

[2]Most of the subjects in this subgroup were able to solve nearly all the problems in category one, and on the average, one problem in four in category two.

they were not sufficiently motivated, and had not taken the trouble to master the propositions. This may be true in a few cases; however, the author can say from personal acquaintance with the group, that a majority of these students were highly motivated--and were very much concerned with their poor performance. One subject, in fact, was constantly seeking help from the author, and although many hours were spent trying to convey the simple ideas to her, progress was practically non-existent. She never seemed to comprehend the necessity of citing evidence for her arguments--triangles were congruent, for example, not because sufficient evidence was at hand to invoke a proposition which proved congruence but rather because "they looked equal", or, "they just were". This kind of 'thinking' would seem to have something in common with what Bartlett (1958) calls 'argument by assertion'.

(g) The examination and analysis of profiles showed that of the 69 students, the upper bounds of 40 could be predicted from the work of the first quarter, of an additional 12 from the second quarter, of 12 from the third quarter, and of the remaining 5 at or beyond the fourth. Thus, the upper bounds of 64 of the 69 seemed to be predictable within the experimental period with sufficient accuracy to form a basis for ability grouping.

<u>Further Profile Analysis</u>

In order to obtain a more specific measurement of the upper bound to use in subsequent calculations, the profiles were subjected to two further analyses.

(a) For a first measure of the upper bound, the 'smoothing' process and determination of the 50% point were carried through as before for each quarter, and for the second half of the term work.

(b) As a second measure of the upper bound, the subject was given credit for the number of categories 'covered' in each quarter. Thus the subject whose profile appeared on page 221 would be said to have covered 1.1 categories in the first quarter and 1.4 categories in the second quarter. Both measures represent an attempt to reduce a profile--which is essentially a multivariate quantity--to a univariate score, and something is undoubtedly lost in the process. However, for practical purposes, either measure is likely adequate to represent the upper bound, since both are sensitive to the relative strength of performance and to changes in performance from quarter to quarter.

Table 38 shows the mean performance for both measures of upper bound over the four categories.

We first notice the general increase between the first and second quarters and the levelling off which is evident

between the third and fourth quarters. This fact, of course, corroborates the individual profile analyses discussed earlier. Since the correlations between the two measures is very high (the scores are, in fact, practically identical), either may be used as a measure of the upper bound.

Table 38

Mean Upper Bound Scores over Four Consecutive Intervals for the Grade 10 Group (N=69)

	50% Point	No. of Categories
First Quarter	1.19	1.17
Second Quarter	1.29	1.30
Third Quarter	1.42	1.41
Fourth Quarter	1.38	1.37

r=.97

The intra-profile reliabilities are of interest since they measure the stability of position relative to the group. Table 39 shows the correlations for the '50%-point" measure.

Table 39

Stability of '50%-point' Measure of Upper Bound for Grade 10 Group (N=69)

First Quarter-Second Half	.70
First Half-Second Half	.84

If the first-half-second-half correlation for '50%-point' scores was used as an estimate of the reliability of the test as a measure of upper bounds, the coefficient, corrected by the Spearman-Brown formula, would amount to .91.

A check on the validity of the '50%-point' as a measure of upper bounds was provided by correlating this score (based on the whole test) against the number of problems solved.[3] For, in the experimental programming, the student was presented with repeated series of problems of increasing difficulty. If the student possessed an upper bound he would tend to reach approximately the same point on the difficulty scale for each set ofproblems. Thus the number of problems solved would be proportional to the upper bound. Since the product moment coefficient of correlation amounted to .96 in this case, it may be assumed then, that the '50%-point' may be used to represent the subject's geometry performance with some confidence in its reliability and validity.

Comparison of Grade 10 and Grade 12 Scores

In the previous section, it appeared that the performance of the vast majority of students reached an upper bound within the experimental period. As a long-range check on

---

[3]The latter score was first converted to a standard score with a mean of 50 and a standard deviation of 10.

this stability, the performance of the matched groups in grades 10 and 12 were compared.

Approximately half of the problems administered to the grade 10 group were also included in the grade 12 battery. This set of problems was culled to remove any that showed temporary practice effects, so that those which remained could be considered of equal absolute difficulty for both groups.

The experienced difficulty (per cent passing) was computed for each of these problems and a comparison appears below.

Table 40

Comparison of Grade 10 and Grade 12 Matched Groups in Geometry Performance

	Grade 12 Superior	Grade 10 Superior	No Difference
First half (21 problems)	14	4	3
Second half (21 problems)	9	10	2

In the first 21 common problems, performance of the grade 12 group was significantly superior (P .05) to that of the grade 10 group; in the last 21 problems the performances were comparable (x^2 test applied in both cases).

The suggestion here, which agrees with the observations of previous sections, is that the majority of grade 10 students attained their upper bounds within a period of approximately 6 weeks (on the average). This is further pointed up in an analysis of the solution of the common problems according to the difficulty categories previously defined.

Table 41

Number of Correct Solutions by Difficulty Categories For the Matched Groups in Grade 10 and Grade 12 Respectively

		Category			
		1	2	3	4
First Half of Term	Grade 10	140	61	72	9
	Grade 12	141	96	130	25
Second Half of Term	Grade 10	118	94	58	15
	Grade 12	122	96	61	22

The performance of the grade 12 group was again superior for categories two, three, and four in the first of the term, but the group performances were comparable over the second half. The slight advantage in favour of the grade 12 group was caused by 5 students who made repeated gains throughout the second year.

It may be plausibly argued, then, that the repetition of a year's work in geometry does not seem to lead to a further increase in the complexity of problems which are correctly analysed by the majority of students. In fact, it appears that most subjects reach an upper bound shortly after they encounter geometry, and that from this time onward, they learn new propositions without increasing the depth to which problems are analysed.

The Effect of Repeating Problems

A third way to test the stability of performance is to study the effect of repeating problems. Most teachers of mathematics believe that a student will profit from being shown how to solve problems which he cannot do by himself. In fact, there is an unwritten law in the teaching profession to the effect that the student must be shown such solutions.

However, the energy theory would not support such a view. If a student's upper bound lies in difficulty category N, then no amount of coaching (outside sheer memorization) would result in his being able to solve problems in category N+L unless an improvement in strategy effects a raising of the upper bound.

This question is examined in Table 42, where two repeated problems in the grade 10 battery and four repeated problems in the grade 12 battery are considered.

Table 42

Increase in the Number of Correct Solutions Upon Repeating Problems

Problem Number	Number of Solutions	Comments
Grade 10		
#1 (15	1	Problems in same exercise.
#1 (20	2	Difficult problem.
#2 (35	26	Problems in different exercises.
#2 (42	18	Problem 42 has higher uncertainty index than 35.
Grade 12		
#1 (9	13	Gain not significant.
#1 (87	17	
#2 (12	4	Significant gain.
#2 (86	22	Additional proposition available for use in the solution.
#3 (54	11	Gain not significant.
#3 (88	16	
#4 (59	6	Gain not significant.
#4 (89	7	

There is no evidence here of a gain due to repetition, except in one case (grade 12, #86). It turned out, however, that the majority of students who obtained a solution on the problem's second occurrence, had available on this later occasion, a proposition which made the solution easier. Sonsequently, the assumption that the demonstration of a problem will increase the probability of its solution is a very tenuous one indeed.

The second pair of the grade 10 repeat problems presents an interesting case; a problem readministered (after the solution was demonstrated) resulted in a decrease in the frequency of correct solutions. This apparent contradiction can be explained in uncertainty terms, however, since the uncertainty difficulty index of the problem was greater in the second probability context than in the first.

Individual Performances

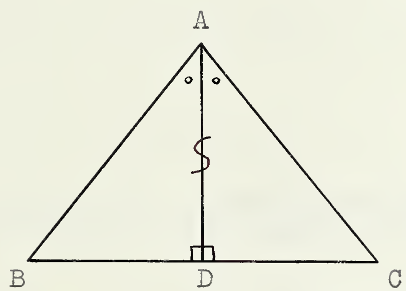
1. Problem-solving behaviour, below, at, and above the upper bound. If the arguments of the previous section are accepted, then one can deduce that the majority of students--using a given strategy and having a fixed maximum quantity of attention-cathexis--will only be able to solve problems up to a certain level of complexity. At this point, the energy supply will become inadequate and the theory has suggested that the highly-motivated and ego-involved student

will ignore alternatives, thus reducing the subjective difficulty of the problem to a level which he can manage.

Perhaps the best way to examine the hypothesis is to consider the solutions offered by a student as he approaches his upper bound and attempts problems beyond this point.

The following 5 problems are part of the battery which followed Proposition XVII. The subject here was a boy with an Otis IQ of 104, and a '50%-point' of 1.2 in geometry. It is fascinating to watch the struggle of this highly motivated student as the problems gradually become too difficulty for him.[4]

(a) Problem One.



Given: Diagram as marked.

Prove: $AB=AC$

Proof offered:

($\triangle ABD$ and $\triangle ADC$
$\angle BAD=\angle DAC$
$\angle ADB=\angle ADC$
AD is common.)

XVII $\longrightarrow$ $\triangle ABD \equiv \triangle ADC$

$\therefore AB=AC$

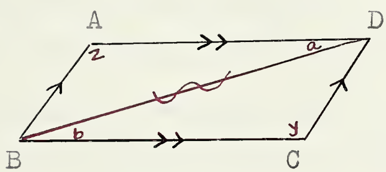
Fig. 39. Geometry problem illustrating performance at difficulty level below the upper bound.

---

[4]The red markings on the diagram were those made by the subject.

Problem one is one of the easiest problems of the first category. The subject's solution is consistent with his expected behaviour, since his '50%-point' rating indicated that he should be able to solve nearly all the problems in this category.

(b) Problem Two.



Given: Diagram as marked.

Prove: AD=BC

Proof offered:

XIV⟶ ∠a=∠b (AD ∥ BC)

XIV⟶ ∠x=∠y (BA ∥ CD)

XVII⟶ △ABD≡△BCD

In △ABD and △BCD

BD is common,
∠b=∠a
∠x=∠y

∴AD=BC

Fig. 40. Geometry problem illustrating performance at difficulty level slightly above upper bound.

This is actually the thirdproblem in the set; the second problem belonged to the first category and was correctly solved.

In problem two, the subject was evidently following the analysis. There is also evidence here of some directed deduction. Realizing that two triangles are needed, the subject has drawn a diagonal and deduced the fact that ∠a=∠b.

Following the analysis table, he apparently knew that he had to use either Proposition II or Proposition XVII. If we attempt to verbalize his subjective elaboration of alternatives it might read as follows:

"To use Proposition II, I would need AB=CD and $\angle$ABD=$\angle$BDC."

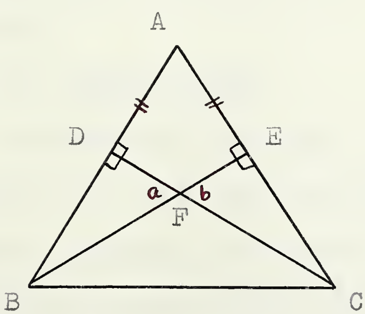
"To use Proposition XVII, I would need $\angle$ADB=$\angle$DBC and either $\angle$BAD=$\angle$BCD or $\angle$ADB=$\angle$DBC."

This uncertainty state has been resolved by reducing the alternatives through what is essentially an assumption until the verbalized solution could be subjectively stated as follows:

"I can use XVII, because I have $\angle$BAD=$\angle$BCD."

Invoking Proposition XIV as an authority is apparently intended to cover the bare assumption with a thin cloak of plausibility.

(c) Problem Three.



Given: AB=AC, and diagram as marked.

Prove: DC=BE

Proof Offered: I $\longrightarrow$ a=b

XVII $\longrightarrow$ $\triangle$DFB$\cong$$\triangle$EFC

In$\angle$DFB and$\triangle$EFC
DB=ED
$\angle$a=$\angle$b
$\angle$D=$\angle$E

$\therefore$DC=BE

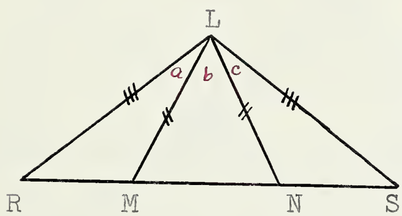
Fig. 41. Geometry problem illustrating performance at difficulty level above upper bound.

This problem, which again belonged to the second difficulty category illustrates even more clearly the subject's logical inadequacy in the face of uncertainty. The fact to be proved (line=line) again called for a choice between Proposition XVII and Proposition II. The subject seemed to favour Proposition XVII, probably because proposition dominance was strong in this set or problems. The next choice involved the selection of triangles. Here, the spontaneous deduction of the fact a=b, apparently led the student into difficulty since it gave the triangles DBF and EFC high subjective probability as the correct choice. Consequently, the other possible pairs--including the correct pair--were not adequately considered. Again the analysis was followed to a point where its subjective complexity exceeded the subject's upper bound. At this point the subject made an uncertainty-reducing assumption which decreased the subjective difficulty of the problem, and thereby rendered it solvable.

(d) Problem Four.

Problem four belonged to difficulty category three. The subject was now well over his head in difficulty and did not really approach an adequate analysis. He identified the type of fact to be proved, and when an elaboration of the possibilities generated a complexity which was completely beyond his power to deal with, he considered the most

immediate propostion which allowed him to prove congruence, viz., Proposition II. Realizing that he needed RLM= SLN, and seeing no way to prove this, he tried to conceal this fact with a weak bluff. Only the faintest outline of an analysis is apparent here. The subject has apparently given up the task of obtaining a completely logical solution, and is merely aiming at plausibility.



Given: Diagram as marked.

Prove: RM=NS

<u>Proof offered</u>:

$\angle a+\angle b=\angle b+\angle c$

$\therefore \angle a=\angle c$

$$\mathrm{II}\longrightarrow \triangle LRM \cong \triangle LNS \left(\begin{array}{ll} \triangle LRM & \triangle LNS \\ \angle R=\angle S & \\ \angle M=\angle N & \\ \angle a=\angle c & \end{array}\right)$$

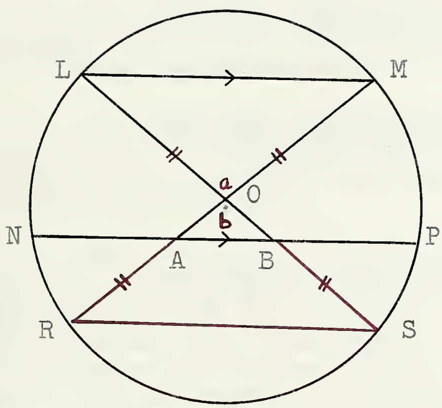
$\therefore RM=NS$

Fig. 42. Geometry problem illustrating performance at difficulty level for above upper bound.

(e) <u>Problem 5</u>.

This problem, which belongs to the fourth difficulty category was so far beyond the subject's capacities that he disregarded the necessity of offering even a plausible solution. In fact, the final statement is in no way connected with the spontaneous deduction that the subject has exhibited. One gets the impression here, of the drowning

man catching at straws. Solutions of this kind become quite incoherent at times, and the subject merely 'sprays' the area with irrelevant statements before giving the required conclusion.



Given: O is the centre of the circle.

: LM || MP

Prove: NA=BP

Proof Offered:

I⟶∠a=∠b

II⟶△LOM=△ROS

△LCM & △ROS

MO=RS

LO=SO

∠a=∠b

∴NA=BP

Fig. 43. Geometry problem illustrating performance at difficulty level far above upper bound.

A complete case history of the subject reveals that his performance in every set of problems showed the same pattern as he approached and passed beyond his upper bound.

Many subjects, of course, merely stop offering solutions when the problems become too difficult. Others develop extremely sophisticated arguments when they find themselves unable to solve problems, attempting to bury an unwarranted assumption in a welter of directionless deduction.

The logical pattern revealed in the above illustrations was found running through the problem-solving behaviour of all subjects who offered solutions for problems above their upper bound. The sequence of logical behavioural stages may be summarized as follows:

(a) Problems below upper bound--Adequate analysis.

(b) Problems near (but above) upper bound--Analysis followed partially. Assumption at some choice point without further consideration of possibilities.

(c) Problems above upper bound--Plausible solution attempted. The alternatives at the first major choice point listed, and sufficient assumptions made to employ the most obvious alternative.

(d) Problems far above upper bound--In the case of a sophisticated subject, the method of attempting to conceal a bluff often took the form of a very lengthy argument in which alternatives were generated to a point where they became extremely numerous and difficult to deal with. At this point some assumption was made.

In the case of a weaker student, the method often began with any obvious deductions which could be made from the diagram. To these were compounded an array of statements--usually quite illogical--concerning the alleged implication of various propositions, and offered in an apparent effort to bedazzle or confuse the reader.

2. The Extent of Illogical Behaviour in Geometry. At this point, it is pertinent to examine the extent of illogical behaviour in the experimental problems. Table 43 partitions the population into groups on the basis of the proportion of correct solutions offered, to incorrect solutions offered.

It should be remembered here, that the numbers

reported are to some extent depended upon the particular battery used. However, since the battery corresponds to the problems ordinarily presented to the grade 10 and grade 12 students in Ontario, it may be claimed that the proportions do not exaggerate the extent of illogical behaviour in the typical classroom. The ratio is distributed so that approximately 65% of the whole populstion offer more correct solutions than incorrect, and 35% offer more incorrect than correct solutions. In this classification, however, only those solutions which were seriously incorrect were considered to lie in the 'wrong' category. When the ratio of correct solutions to the sum of wrong-plus-partially-correct solutions was computed and tabulated, the distributions as shown in Table 44 resulted.

In this case, only 46% of the population offered more correct than incorrect solutions. The claim that geometry affords the student an opportunity to practice and develop systematic and logical methods of analysis should be soberly weighed in the light of actual classroom conditions. Indeed, for a majority of students in the normal classroom, geometry would seem to provide more exercise in the use of illogic than in the use of logic. Under these circumstances, it is difficult to imagine that any positive transfer will take place when the student attacks problems in real life.

Table 43

Distribution of The Ratio: (Correct Solutions) to (Incorrect Solutions), for Grades 10 and 12

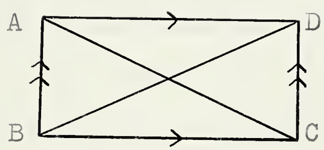
	Above 10-1	5-1 to 10-1	2-1 to 5-1	1-1 to 2-1	1-1 to 1-2	1-2 to 1-5	1-5 to 1-10
Grade 10	5	8	16	15	12	11	2
Grade 12	9	5	5	6	5	5	2
Total	69 (65%)				37 (35%)		

Table 44

Distribution of the Ratio: (Correct Solutions) to (Wrong and Partially Correct Solutions), for Grades 10 and 12

	5-1 to 10-1	2-1 to 5-1	1-1 to 2-1	1-1 to 1-2	1-2 to 1-5	1-5 to 1-10
Grade 10	3	10	18	16	20	2
Grade 12	1	13	4	8	9	2
Total	49 (46%)			57 (54%)		

3. Effect of Expectation of Solution. A rather interesting result turned up when some information concerning a diagram used in the solution of a problem was accidentally omitted. The incompletely-stated problem read as follows:



Given: ABCD is a parallelogram.

Prove: $\angle A = 90^\circ$

Fig. 44. Geometry problem illustrating the effect of an expected solution.

The result, of course, simply cannot be proved from the given information. Nevertheless, 21 of the 37 grade 12 students presented an argument which concluded with the required fact.

The geometry student is nurtured on problems, nearly all of which are solvable. In time, he comes to develop a high expectation that the problem which confronts him at the moment will be solvable, provided that the proper analysis is made. When he encounters one which looks relatively easy, but which is, in fact, incapable of solution, he often believes that he has overlooked some simple fact. If he fails to find a solution, and if he is sufficiently ego-involved to see an ego threat in his inability to find one, then he is only a short step away from inventing a plausible-

looking solution.

Summary

Three sources of evidence offered support for the 'boundedness' hypothesis of the energy theory. In the first place, the majority of students (79%) reached their 'natural' level of performance within the first three weeks, and, from that time onward did not make sufficient gains to have moved into a higher category by the end of 11 weeks. The performance of all but 5 of the 69 grade 10 students exhibited upper bounds within the experimental period.

A second source of confirmation came from studying the increase in the frequency of correct solutions when problems were explained to students, and then repeated. There was no evidence of a significant gain here, and it appeared that if the problem difficulty is above the student's upper bound, an explanation in itself will not extend the upper bound to the point that the student will be able to solve the problem at its second occurrence.

The third confirmation came from comparing the frequency of solutions of problems common to the batteries administered to the matched grades 10 and 12 groups. The performance of the grade 12 group was significantly superior on the problems which occurred during the first 5 weeks, but once the grade 10 students reached their optimum level, the

performance of the two groups was comparable.

In summary, it would seem that approximately half of the grade 10 population was not able to proceed much beyond the immediate applications of the propositions. Moreover, for those students who continue to grade 12, it seems (for the vast majority of students) that all that a second year of geometry accomplishes is to add a number of reference points without increasing the students' power of analysis.

Much more alarming than the general inability of a large segment of the student population to make much progress in geometry, is the type of behaviour exhibited when the student encounters a problem which is too difficult for him. For, while the inability to solve a problem is regrettable, the willingness to force a solution to fit a preconceived conclusion is surely a dangerous mental habit to encourage in the student. And yet, if our population is a representative one, it would seem that a majority of students spend the greater part of their time illogically reducing problems which are too difficult for them.

It would be unfortunate if this attitude carried over into the students' everyday life, where the complexity of the problems encountered is usually so vast that immediate solutions are simply not available. It would likely be to the advantage of the student, if he were able to suspend

judgment in the face of problems which are too difficult for him to deal with adequately. However, our school subjects--and geometry in particular--are strongly 'solution' oriented. The student is never rewarded for refusing to offer a solution; in fact, partly-correct solutions usually earn the student marks on an examination. It may be that the day-by-day t eaching of geometry systematically destroys the very attitude of impartial inquiry which it claims to inculcate.

CHAPTER XIII

PROBLEM SOLVING UNDER INFINITE UNCERTAINTY

The previous study dealt with problems in which the alternatives were finite in number. However, when the matter is considered at any length, it becomes apparent that such situations form the exception rather than the rule--for problems in real life rarely present a well-defined, enumerable set of alternatives; thus, the first illusion of a limited number of alternatives invariably gives way to a vast complexity of factors.

In this section, we were interested in the extent of strategy formation in infinite uncertainty situations, and in the extent of guessing and simplification which occurs at this level.

Test Battery, Subjects, and Procedure

Considerable thought was given to the construction of a battery of problems which would satisfy the following requirements:

(a) The problems should generate a potentially infinite set of alternatives.

(b) The battery should consist of problems of increasing difficulty, so that a measure of upper bounds is present.

(c) The problems--at least at the lower levels--should be numerous enough to allow a determination of reliability.

(d) The problems should be systematically varied so that the principle used to solve one would not necessarily solve its successor.

(e) Some consideration should be given to the location of problems so that the practice effects, if any, could be assessed.

The author employed the previously mentioned number series which were constructed by the principle of finite differences.[1] These series can be considered as instances of problem solving under infinite uncertainty, for a reason which will become apparent upon examining a specific series. For example, when the student encounters the series:

2, 5, 13, 22, 84, ___.

he will undoubtedly begin by considering the relationship between the first two numbers. However, it is possible to formulate an infinite number of possible relationships between them; thus 5 is three more than two, 5 is three times two less one, 5 is two squared plus one, and so on. One or more of the possible relationships is tested on the second pair of numbers, and those surviving, on the third, and so on. The chief reason why the individual is able to solve such series at all is that he seems to possess a preferred order of testing hypotheses, i.e., certain alternatives have high subjective probability.

---

[1]The investigation presently considered was carried out prior to the use of the number series in the group experiments.

Number series with differences of orders one, two, three, and four were embedded in a set of miscellaneous number series to yield the following battery.

Table 45

Problem Battery[2]--Infinite Uncertainty

Practice

A - First-order difference (increasing).
B - Second-order difference.
C - Pattern.
D - First-order difference (decreasing).

Set A	Set B
1. First-order (increasing).	1. First-order (decreasing).
2. Pattern.	2. Pattern.
3. Second-order.	3. Second-order.
4. Multiplication.	4. Division.
5. First-order (increasing).	5. First-order (increasing).
6. Rhythm.	6. Rhythm.
7. Third-order.	7. Third-order.
8. First-order (decreasing)	8. First-order (increasing).
9. Division.	9. Multiplication.
10. Second-order.	10. Second-order.
11. Fourth-order.[a]	11. Fourth-order.

[a]Not used unless subject solved third-order series.

In the multiplication series, each successive member of the series is obtained by multiplying its predecessor by a fixed constant. The rhythm series of the type,

5, 10, 15, ___.,

[2]The actual problems may be found in Appendix E.

is, in fact, a first-order difference series, but it is of the type that children chant when 'counting by 5's'. It seemed likely that such series would be recognized in less time than other first-order series because of this familiarity.

The problems were typed on individual cards which were presented ot the subject one at a time. The time for each solution was recorded with a stop-watch, together with the correct or incorrect solution.

Before the testing session began, the following directions were read to the subject:

> On these cards you will find a set of numbers, followed by a blank. Your job will be to look at the numbers and see if you can figure out what number should go in the blank. You will find some of the questions to be very easy, but some will be difficult and will require a considerable amount of time to solve. In any case, you may take as long as you please to do them. Remember that accurate solutions are more valuable than fast solutions. The first four are practice questions.

At this point, the four practice cards were presented. If the subject solved any problem incorrectly, or if he was unable to solve it, then help was given, but no general method of solution was indicated. In the battery proper, no help of any kind was given, nor was any writing or computing allowed.

Results

1. Reliability. Median values of the three first-

order series scores on each half of the test proper were correlated, yielding the reliabilities shown in the following table.

Table 46

Reliability Estimates for First-Order Number Series

	Product-Moment r	Corrected by Spearman-Brown Formula
Grade 10 (N=69)	.70	.83
Grade 12 (N=37)	.86	.93

The high reliabilities are surprising for two reasons. In the first place, normal problem solving itself seems to be a somewhat erratic process and 'discovery' times tend to vary. Again, the total duration of the test (6 questions) was approximately 45 seconds and 36 seconds for grades 10 and 12 respectively. It would seem that the subjects, and this is especially true of grade 12, tended to attack the problems in a highly consistent manner. In fact, a product-moment correlation of .77 (corrected to .87) existed between the grade 12 times on problem A-#1 and A-#5, although the latter problem was considerably more difficult.

2. Distribution of Upper Bounds. The upper bounds were readily determined in this case by taking the highest

level at which the subject solved one or more problems correctly. The results for both grades are shown in Table 47, and are there compared with the upper bounds which were exhibited in the group test in Level I performance (Chapter VII).

Table 47

Upper Bounds in Problem Solving Under Infinite Uncertainty (Level II); a Comparison with Upper Bounds in Level I Performance

		Order				
		1	2	3	4	5
Grade 10	Series Level II	41	25	1	0	0
N=67[a]	Series Level I	7	24	21	10	5
Grade 12	Series Level II	10	21	5	0	0
N=36[a]	Series Level I	0	8	13	9	6
Total	Series Level II	51	46	6	0	0
N=103	Series Level I	7	32	34	19	11

[a]Three cases were deleted from the Level II Series which corresponded to the missing cases in the Level I Series data. The contradicting cases in the Level I data were recorded as the lowest level which could be interpreted as the upper bound.

It must be remembered, of course, that the Level I test was not really a set of problems, since the rules for solution had been provided. Nevertheless, this latter score provided a minimum estimate of the student's potential

problem-solving level; minimum, because we saw that this level could be raised if group pressure was reduced. In other words, it provided us with an estimate of what the student was able to do in terms of the operation-storage schema, provided that he knew how to proceed. For this reason, it was a convenient measure against which problem solving could be compared.

The first thing which impresses one on looking at this table, is the outstanding lack of problem solving exhibited. Almost half of the total population solved no problems beyond the first order. Moreover, there was no evidence of a strategy development during the completion of the battery. For example, the number of correct solutions for the first second-order difference problem (A-#3) amounted to 31, while the correct solutions for the last second-order difference problem (B-#10) numbered 28.

One may legitimately ask whether a student can be expected to develop a strategy for dealing with a set of alternatives which is potentially infinite. Indeed, the energy theory would predict little progress when attention is divided among a large number of alternatives. What is of more interest here, is the type of strategy employed by the student who consistently obtains a false solution. The following sections consider the matter at greater length.

3. Analysis of Errors at Higher Levels. In Tables 48 and 49, the results of the analysis of individual solutions are exhibited under the categories listed.

Several things emerged from these tables. In the first place, a significantly greater proportion of second-order problems were solved as first-order problems than as second-order problems by the grade 10 group.[3] At levels three and four, the proportion of problems solved as types of a lower order was again significantly greater than the proportion solved correctly.

In the grade 12 group, the proportion of correct solutions was larger for the second-order (not significant), but the proportion of problems solved as a lower order type for orders three and four was significantly greater than the proportion solved correctly.

One alarming result was that the proportion of students who offered no solution did not show an increase as the problem difficulty level rose. This was in sharp contrast to the group tests, in which there was a significant increase in the proportion of students offering no solution as the difficulty level rose. We shall interpret this to

---

[3] The sign test was employed in which the subject's performance was marked '+' or '-' depending on whether he solved the larger proportion of second-order problems as second-order or as first-order questions. Significance was tested at the .05 level in each case.

Table 48

Analysis of Solutions for Level II Number Series--Grade 10

	Order		
	2	3	4
Solved Correctly	44	1	0
Solved as First-Order	147	75	22
Solved as Second or Third-Order	-	-	4
No Solution Offered	11	18	7
Response Not Classifiable	69	37	34

Table 49

Analysis of Solutions for Level II Number Series--Grade 12

	Order		
	2	3	4
Solved Correctly	60	8	0
Solved as First-Order	45	29	13
Solved as Second or Third-Order	-	9	5
No Solution Offered	7	10	9
Response Not Classifiable	32	16	9

mean that whereas the student who had the rules, guessed because he was under pressure to do so, the problem solver deluded himself concerning the correctness of his solution. Expressed otherwise, it would indicate the relative ease of self-deception which comes from a systematic ignoring of part of the given data. It also indicates that the subjects were willing to attack problems far beyond their comprehension by bringing this systematic over simplification into play.

Factors Influencing Hypothesis Formation

The number series presents an infinite array of alternatives. Within that array, the student seemed to make a finite selection. In this section, we will consider some factors which influenced this selection.

The battery contained 8 first-order series mixed in with the general battery so that one principle would not suffice to solve consecutive items. The series were originally chosen to be of equal difficulty. Once the choice of series was made, they were randomly assigned to the 8 possible positions.

It seemed to be a reasonable hypothesis that the alternatives which would occur to and be tested by the student, would depend upon two sources: (a) habitual strategies which expressed themselves in constant modes of approach to the series, (b) temporary 'sets' (subjective

probabilities to favour and test a particular alternative) induced by the structure of the test battery. Under these circumstances four factors would seem to be involved in the experienced difficulty of a series: (1) long-range subjective probabilities favouring particular alternatives, (2) probabilities induced by the position of the series within the battery, (3) the inherent difficulty of the question, (4) the effect of practice on computing speed.

1. Practice Effects. Since some of the subsequent argument will assume no appreciable practice effect, this hypothesis will be checked at the outset. Table 50 lists the mean times in seconds for the first and last three problems for both grades.

Table 50

Practice Effects in First Order Number Series

	Mean Time Per Problem For First Three Problems	Mean Time Per Problem For Last Three Problems
Grade 10	7.33 sec.	7.53 sec.
Grade 12	5.94 sec.	6.03 sec.

An examination of the problems in question shows that both sets of three contain two increasing, and one decreasing series, and that the numbers involved (and the increments)

problems were randomly assigned to their respective cells, the probability of the observed agreement in the order of difficulty would be very small, much less than .01 in fact, so we may dismiss the hypothesis that positional effects are not present.[4]

Table 51

The Effect of Temporary 'Sets' on the Solution of First-Order Problems

	Gradient of Decreasing Similarity Time in Seconds				
Type of Problem Following	D_1	D_2	D_3	Mult.and Div.	Pattern
Grade 10, N=69	5.92	6.39	7.91	8.73	12.1
Grade 12, N=37	4.30	5.21	6.04	7.16	8.95
Difficulty Rank For The Whole Group	1	2	3	4	5

3. <u>The Effect of Long-Range Strategies</u>. Over and above momentary sets operate certain long-range strategies. These ingrained modes of dealing with sequences of numbers produce their effect in the delimitation of alternatives.

---

[4]The phenomenon of 'set' seen here lends further plausibility to the assumptions underlying 'proposition dominance'.

problems were randomly assigned to their respective cells, the probability of the observed agreement in the order of difficulty would be very small, much less than .01 in fact, so we may dismiss the hypothesis that positional effects are not present.[4]

Table 51

The Effect of Temporary 'Sets' on the Solution of First-Order Problems

	Gradient of Decreasing Similarity Time in Seconds				
Type of Problem Following	D_1	D_2	D_3	Mult. and Div.	Pattern
Grade 10, N=69	5.92	6.39	7.91	8.73	12.1
Grade 12, N=37	4.30	5.21	6.04	7.16	8.95
Difficulty Rank For The Whole Group	1	2	3	4	5

3. <u>The Effect of Long-Range Strategies</u>. Over and above momentary sets operate certain long-range strategies. These ingrained modes of dealing with sequences of numbers produce their effect in the delimitation of alternatives.

[4]The phenomenon of 'set' seen here lends further plausibility to the assumptions underlying 'proposition dominance'.

The battery was constructed so that differences in the speed of solution could be compared for different types of series. The first line in Table 52 presents the postulated order of difficulty on the hypothesis that the subject, when presented with a card, would first respond to the general pattern or configuration of numbers, and that if no pattern existed, he would test the hypotheses in the order of their highest subjective probability. Since the subjective probability would likely depend upon the frequency of occurrence in the individual's past experience, this would mean that addition, subtraction, multiplication and division hypotheses would be tested in that order. The experienced order of difficulty is given in the second line of Table 52.

Table 52

Postulated and Experienced Order of Difficulty for Level II Number Series

	1	2	3	4	5	6	7
Postulated Order	Pattern	Rhythm	D_1	Mult.	Div.	D_2	D_3
Experienced Order (Whole Population)	Rhythm	D_1	Pattern	Div.	Mult.	D_2	D_3

A series of t-tests yielded the following results for both the grade 10 and grade 12 groups.

(a) The rhythm pattern required significantly less time (P .01) that all the others.

(b) The D_2 and D_3 series requires significantly greater time (P .01) than all the other series.

(c) The differences between solution times for successive steps in the sequence: D_1, Pattern, Division, and Multiplication, series were not significant, but the Addition-Subtraction series (D_1) required significantly less time for solution than the Multiplication-Division series (P .01).

There is general agreement, here, with the notion that the subjective probability of an alternative, and therefore its order of testing, is determined by the frequency of its occurrence in the past experience of the subject.[5] Thus most subjects had such high expectancy for difference series that they could recognize and solve these problems faster than a simple pattern such as

1, 2, 2, 1, 2, 2, ___.

Matched Groups

The matched groups at the grade 10 and grade 12 levels were compared as in the previous experiment. The

---

[5]This conclusion is subject to the reservation that it was not entirely possible to take out or remove the effect of temporary sets in this last analysis, and it is possible that such effects acted differentially on the various types of questions.

mean D_1 and median D_1 scores are presented in Table 53.

Table 53

Mean and Median D_1 Scores for Grades 10 and 12 Matched Groups

	Mean of Average Times in Seconds	Mean of Median Times in Seconds
Grade 10	7.27	6.43
Grade 12	6.41	5.81

In both cases, a comparison of means showed the grade 12 group to be faster, but in both cases the differences just failed to reach significance at the .05 level. Combining this result with Table 51, it would seem to be a safe inference that the grade 12 student has a better-defined strategy, and that his strategy is less affected by temporary 'sets'. The large discrepancy between mean and median scores was caused by the fact that some subjects seem to 'block' on one or two questions, thus raising the mean without seriously affecting the median.

A second source of comparison involved the frequency of solution of problems above first-order differences. An examination of upper bounds (Table 54) indicated the superiority of the grade 12 group.

Table 55 contrasts the number of problems of orders

Table 54

Comparison of Upper Bounds in Level II Series for Grade's 10 and 12 Matched Groups

	Order			
	1	2	3	4
Grade 10	15	14	1	0
Grade 12	9	18	3	0

Table 55

Number of Correct and Incorrect Solutions to Level Number Series Above Order 1--Matched Grade 10 and 12 Groups

	Grade 10	Grade 12
Solved Correctly	26	47
Solved as First Order	99	80
% Correctly Solved	20.8	37.0

two and above which are solved correctly with the number solved as first-order problems.

While neither result was significant at the .05 level, the combined evidence, together with the first-order results indicated the superiority of the grade 12 group beyond reasonable doubt.

This would indicate that the grade 12 group was either influenced less by temporary sets, or that the additional two years of experience with numbers has altered their permanent subjective probabilities concerning alternatives in number series. The latter hypothesis is made plausible by the fact that the intervening grade 11 algebra course contained a section on arithmetical (D_1) series and geometrical (M) series.

Summary

Many social and personal issues with which the student must eventually come to grips are so complex as to present almost infinite uncertainty. In this section, the logical behaviour of a group of students in one infinite uncertainty situation was studied. No evidence of positive strategy formation during the test period was found. Rather, the student seemed to bring into the testing situation a set of well-defined subjective probabilities which delimited and established an order of precedence among alternatives.

These subjective probabilities were clearly related to past experience, for the high expected frequency of 'difference' series resulted in their being recognized in less time than simple patterns of numbers.

A great deal of negative strategy formation was found, however. More problems of the second order and above were solved as problems of the first order than were solved correctly. This was again partly attributable to temporary sets introduced within the battery, and partly to the permanent first-order expectations. The regretable thing here, is again, not so much that the subject could not solve the problems, but that he was willing to offer solutions which were clearly inconsistent with the complete data presented. What this means in effect is that if a subject has a strong preconceived notion, he will conveniently ignore that part of the data which contradicts his limited hypothesis.

CHAPTER XIV

CORRELATES OF PROBLEM-SOLVING ABILITY

Individual Predictors of Problem Solving Ability in Geometry

It was seen in the geometry study that problem-solving ability could be accurately assessed by the end of the first three weeks. It would, of course, be more useful to be able to make a completely a priori prediction. This section examines the extent to which this is possible.

The standard battery of tests which was administered at the beginning of the experimental period, provided a set of scores which could be interpreted, in the case of grade 10, as predictors, and in the case of grade 12, as correlates, of geometry problem solving ability.[1]

Several observations may be made in connection with Table 56.

The first unusual finding was that Otis IQ ranked second among the single predictors at the grade 10 level, and third among the single correlates of geometry ability at the grade 12 level. A further analysis of individual cases

[1]The high correlates of ability in geometry at the grade 10 level need not necessarily be good predictors of geometry ability in general. We would leave the word 'predictor' to refer to a variable whose value may be determined prior to the behaviour to be predicted. The covariation of certain variables with geometry ability at the grade 12 level can be partly attributed to the selection methods used in the school.

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

Table 56

Single Predictors (Grade 10) and Single Correlates (Grade 12) of Problem-Solving Ability in Geometry[a]

	Grade 10	Grade 12
1. Chronological Age	-.19	-.11
2. D.A.T. Numerical	.40	.66
3. D.A.T. Spatial	.48	.71
4. K_1	-.23	-.35
5. K_2	-.02	-.09
6. Minn Spatial	.32	.42
7. Otis	.42	.57
8. Study Habits	.08	.24
9. Watson-Glaser	.25	.55

$r_{p=.01}$ for grade 10 is .31, $r_{p=.05}$ is .23

$r_{p=.01}$ for grade 12 is .42, $r_{p=.05}$ is .31

[a]The measure of problem-solving ability employed here and throughout was the 50%-point.

revealed that at the grade 10 level particularly, many superior problem solvers had mediocre Otis IQ's; conversely, some students with very superior IQ's were very poor problem solvers. The rise in the correlation for grade 10 to grade 12 can probably be explained by the fact that the low IQ-high problem solver may never reach the academic course in grade 12, since verbal limitations mean that his school marks will be generally low.

The grade 12 correlations between the standard tests and the geometry scores were all higher than those for grade 10. The small number of subjects in the grade 12 (sample N=37) precluded the possibility of establishing significance for individual coefficients, but the whole array of coefficients is significantly higher for grade 12 than for grade 10. Again, this probably means that the high-ability, low-motivation student who lowers the correlation coefficient drops out before grade 12.

The large difference in predictive value of the Watson-Glaser between the two grade levels requires special attention. The author suspects that this test may have been too difficult for many students of the grade 10 group despite the fact that the test manual indicated standardization down to the grade 9 level.

The high correlations with the D.A.T. Spatial at both levels is somewhat surprising. The contribution of spatial

ability to geometry performance has long been argued and it cannot be said to have been settled yet. There is nothing in the plane geometry diagram which would seem to require the ability to visualize how an object would appear if rotated in space, or the ability to visualize the three-dimensional shape of an object which would result from folding a given two-dimensional form.

Both geometry problems and the spatial test require extreme and continued attention, however, and in both, the subject is heavily penalized for incautious behaviour. The high IQ-low spatial subjects usually showed a surprising number of incorrect answers, both in the D.A.T. Spatial and in problem solving.

It is apparent that little relationship existed between geometry problem solving ability and either K_1 or K_2. This fact seems to agree with the theory, which sees comprehension as a psychological function, distinct from problem solving.

In summary, it may be said that while there seemed to be several sizable correlates of geometry at the grade 12 level, no individual prediction variable at the grade 10 level was sufficiently large to allow accurate prediction.

Multivariate Prediction of Ability in Geometry

Only 5 of the original predictors showed correlations

with geometry which were significant beyond the .05 level. Since the two spatial tests were highly intercorrelated, the Minnesota was dropped, leaving the following four predictors:

(1) D.A.T. Numerical

(2) D.A.T. Spatial

(3) Otis

(4) Watson-Glaser

In Table 57 and Table 58, the intercorrelations between these variables have been listed. The coefficients were determined by the Doolittle method[2] (McNemar, p. 183), and they, together with the multiple coefficient of correlation, appear in Table 59.

Table 57

Intercorrelations of Four Predictors[a] of Geometry Ability Grade 10

	1	2	3	4	Geom.
1		.12	.45	.25	.40
2			.24	.25	.48
3				.64	.42
4					.25

Code: 1. D.A.T. Numerical 2. D.A.T. Spatial 3. Otis 4. Watson-Glaser

[2]The assumption of linearity of regression was checked by inspection of the scattergrams.

Table 58

Intercorrelations of Four Correlates[2] of Geometry Ability
Grade 12

	1	2	3	4	Geom.
1		.66	.51	.40	.66
2			.48	.23	.71
3				.52	.57
4					.55

Code: 1. D.A.T. Numerical 3. Otis
2. D.A.T. Spatial 4. Watson-Glaser

Table 59

β Coefficients and Multiple R for Grades 10 and 12

	Grade 10	Grade 12
1. D.A.T. Numerical	.282	.173
2. D.A.T. Spatial	.408	.476
3. Otis	.253	.080
4. Watson-Glaser	-.076	.331
$R_{5.1234}$	.63	.82
$R^2_{5.1234}$	.396	.680

[illegible]

[illegible]

[illegible]

[illegible]

Again, the D.A.T. Spatial emerges as the best predictor, while the Otis is relegated to the third and fourth positions of importance in grades 10 and 12 respectively.

The combined predictors explain approximately 40% of the variance in problem solving scores at the grade 10 level, and are clearly inadequate for accurate prediction at the group, let alone the individual, level. It would seem, then, that it may be impossible to predict the progress that the student will make in geometry problem solving without actually sampling some of this behaviour.

Intercorrelations of Strategy Formation

Level I mathematical ability emerged as a distinct cognitive function--the two tests in this area showed high intercorrelations. It is of interest to enquire to what extent the measures of problem solving are related. More particularly, we are interested in finding out whether strategy formation is a general factor which runs through the several tests, or whether it is specific to a particular test.

Four measures of strategy formation were used: (1) gaps, (2) Level II series, (3) ordering, and (4) geometry. The method of deriving the gap score has been previously discussed. A similar scale was set up for the ordering problems. The series score was t aken as the subject's

exhibited strategy and was evaluated on the basis of whether he did (+) or did not (-), solve correctly, problems at, or above second order differences. The geometry limit score was taken as a measure of strategy formation for that test.

The results of intercorrelating the four tests are shown in Table 60 and 61.

The significant intercorrelations for the grade 12 group suggests the presence of a common strategy-formation factor running through them. This is especially noteworthy when the diversity of situations is considered. Three of the grade 10 coefficients were significant and three were positive but not significant. The generally low intercorrelations have undoubtedly been influenced by the fact that the grade 10 group exhibited little strategy formation on any test.

If we probe beyond strategy formation for underlying personality traits, we immediately find an inverse relationship between the tendency to guess and the amount of strategy formation exhibited. This is natural enough, because the individual who does not form an adequate strategy, must either guess or refuse to offer a solution. The product moment coefficient of correlation between 'guessing' socres and limit scores in geometry amounted to -.55.[3]

---

[3]The guessing score in geometry was determined from

[illegible]

Table 60

Intercorrelations of Four Measures of Strategy Formation
Grade 10

	Gaps	Level II Series	Ordering
Geometry	r_{bis}=.40**	r_{bis}=.16	r_{bis}=.32*
Gaps		ϕ =.17	ϕ =.53**
Series			ϕ =.19

Table 61

Intercorrelations of Four Measures of Strategy Formation
Grade 12

	Gaps	Level II Series	Ordering
Geometry	r_{bis}=.45**	r_{bis}=.50**	r_{bis}=.47**
Gaps		ϕ =.40*	ϕ =.51**
Series			ϕ =.36*

* Significant at .05 level.

** Significant at .01 level.

[illegible]

[illegible]

[illegible]	[illegible]	[illegible]	
[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]		[illegible]
[illegible]			[illegible]

[illegible]

[illegible]

[illegible]	[illegible]	[illegible]	
[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]		[illegible]
[illegible]			[illegible]

[illegible]

[illegible]

The tendency to guess--or the lack of caution--is characteristic of the poor problem solver in every area. It should follow then, that significant correlations would be found between indices of guessing in different situations. The hypothesis may be investigated by examining Table 62.

It turned out that only one of the three indices was significant, although the other two were positive and approaching significance at the .05 level. It is interesting to note the relatively high correlation between the two indices of guessing in group behaviour and the relatively low correlation between indices from group and individual situations.

It would seem, then, that the individual develops a stable guessing pattern which is a function of his position in terms of performance relative to the class, and which runs through both problem solving and comprehension behaviour. The consistency of the index depends on the fact that he

---

the formula,

$$G=\frac{N_2}{N_1+N_2}$$

where N_2=number of problems solved incorrectly,

N_1=number of problems attempted, but no solution offered.

The coefficient G thus represents the relative proportion of problems that the subject is unable to do, for which he subsequently offers an incorrect solution.

tends to find himself in the same position relative to the class over long periods of time.

Table 62

Intercorrelations of Three Indices of Guessing[4]--Grade 12

	Level I Series	Level II Series
Geometry	$r_{bis}=.54^{**}$	$r_{bis}=.32$
Level I Series		$\phi=.26$

** Significant at .01 level.

Apparently the guessing patterns for the grade 10 students had not yet crystallized. This was perhaps due to the fact that they were in a newly formed group and had not yet clearly perceived their position within the group. In any case, the three coefficients for grade 10 were all positive and of the order of .3 for the group-group correlations and of .2 for the group-individual correlation.

---

[4]The guessing score (dichotomy) was determined as follows:

(a) Level I Series: guessed (-) or did not guess (+) at stage 5.

(b) Level II Series: solved second-order problems as first-order (-) or as second-order (+).

[illegible]

Sex Differences

The question of sex differences is naturally and invariably raised in any study of mathematical ability. There is a common belief that boys are superior to girls, especially with respect to ability in geometry. The data provided an opportunity to examine such differences.

The general finding was that for the sample employed in this investigation, no substantial or statistically significant difference was found in mean performance scores between the sexes in any of the standard or experimental tests, or in geometry performance.

Slight differences were observed in favour of the boys in the D.A.T. Spatial, the Minnesota Form Board, the Watson-Glaser, and in the geometry limit scores. Small differences favouring the girls existed in the Otis IQ, the Dominion IQ, the Brown-Holtzman Study Habit Inventory. Scores on all other tests were nearly identical.

It has often been found, and is a commonly accepted fact that in the general population the boys' performance on spatial tests is markedly superior to the girls'. The fact that this did not prove true in our sample, coupled with the slight superiority of the girls' IQ, might suggest that we actually had a relatively inferior group of boys. This in turn, might be argued as a possible explanation of

the lack of sex difference in geometry performance.

Although this question cannot be settled here, the author doubts the importance of 'spatial' ability--as measured by the tests employed in this study--as a factor in success in plane geometry.

What would seem to be important--and some evidence can be cited--is the part played by the sensitivity to different sex roles, and their relationship to geometry. A recent study by Kelland (1959) indicated that in our secondary schools, the boys achieved the greatest status--in the eyes of teachers, parents, and pupils--by high performance in mathematical and scientific subjects, while the girls achieved status through high performance in verbal subjects.

Two facts tended to support this argument in the present study. The following, rather provocative question, was put to the students at both levels (Questionnaire 1, #11):

> "Do you think that boys should do better than girls (of equal intelligence) in mathematics?"

The interesting result is that while girls at the grade 10 level (especially those who perform well in mathematics) invariably answered, "No", to the question, girls at the grade 12 level were more evenly divided on the issue. In other words, it seems that the grade 10 girl has not singled out areas in which boys are expected to excel, whereas grade

12 girls have differentiated such areas.

Another interesting fact was found in the score on the conformity scale of the Minnesota Counseling Inventory. For the boys, no clear pattern emerged in the relationship between conformity and geometry performance at either level. For the girls, however, high geometry performance correlated with relatively high conformity at the grade 10 level, and with relatively low conformity at the grade 12 level. There is also an indication--although it is not statistically significant--that high performance in geometry among grade 12 girls is associated with a low social-relationships score.

There is some indication here, that the girl who performs well in a 'male' area, tends to flout her prescribed role to some extent, and may suffer socially for doing so.

Rural and Town Students

The population was fairly evenly divided among rural and town students. Contrary to expectations, the rural students held a superiority in mean geometry scores at the grade 12 level (significant at .05 level), while the difference in mean IQ was not significant. Differences at the grade 10 level were not significant. Since the finding seems somewhat contrary to accepted doctrine, it was examined in greater detail.

An examination of other test scores revealed that the

town group was superior in the so-called skill subjects, arithmetical computation being a prime example. On the other hand, the rural group seemed to be better at arithmetical and geometrical problem solving. It may be that the rural student, who is left to work out his problems more-or-less by himself, is, in the long run, provided with a better training in problem solving, which must be, in fact, a solitary process. Comparative studies of achievement rarely include adequate measures of problem-solving ability.

It might be possible to interpret the rural student's superiority in geometry at the grade 12 level in terms of superior motivation. One might expect to find apathy toward school among some farm and twon students at the grade 10 level; many students are, in fact, 'waiting out' the legal school leaving-age requirement. However, when the apathetic rural student reaches age 16, he invariably leaves school--on the one hand because his parents are inclined to adopt the 'go-until-you-fail' policy, and again because a readily available and steady source of occupation is ready to absorb him.

The apathetic town student, on the other hand, is eligible only for the most menial sort of labour, and so he is often found in the higher grades of school.

Expectation Toward Mathematics

It is often contended that a student's performance in mathematics may be influenced--beneficially or adversely--by the expectations that he carried with him when be begins the mathematics course. These expectations probably emanate from two main sources: (a) early experiences in mathematics, (b) experiences of parents or siblings in mathematics.

Dealing with (a) first; by far the greater part of our population (64%) found public school arithmetic to be one of their easier subjects, while only 29% put geometry in the same category. Moreover, there was strong agreement (70%) in declaring geometry to be harder than algebra. Eighty-five per cent of the superior geometry students (top half) and 44% of the weaker students found arithmetic fairly easy. Thus we may say that almost all the students who were later to perform well in geometry had early expectations of success, while about half of the students who were later to do poorly, had early expectations of failure.

These observations are supported by answers to the question (Questionnaire 1, #4),

"When did you first become aware that mathematics was going to be easy (or difficult) for you?"

Again, a majority (51%) had made this decision in public school (grade 6 was mentioned most frequently). The beginning of algebra convinced another 19%, while the final

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

30% did not seem to decide until they encountered geometry. There was striking indication from the answers that many students felt that they could handle the computational and comprehension parts of arithmetic and algebra, but that they had found problem solving difficult--as early as grade six.

With regard to category (b), most of the students who performed well in geometry (88%) claimed that their parents also found it easy, while 68% of the weak students claimed that their parents had also found mathematics difficult. Several extreme cases of overachievement (i.e., mediocre IQ and high geometric performance) could be partially linked to expectations derived from parents. For example, the highest performing grade 10 student--a girl with an Otis IQ of 102--when questioned about her performance in geometry, said that she knew geometry would be easy for her because her mother had always stood at the top of the class in mathematics.

CHAPTER XV

THE PROBLEM-SOLVING SYNDROME

Problem-solving ability, as defined in this thesis, appears to be a fairly general cognitive factor, which shows only moderate correlation with Otis IQ, and which seems to have no outstanding predictors in the cognitive area. In an attempt to find non-cognitive correlates of this ability, the author correlated standardized geometry limit scores against a variety of personality and socio-economic variables.

The results were meagre and rather discouraging--the chief results have already been discussed under sex differences, expectations toward mathematics, and rural and town students.

The study of overachievers in geometry turned out to be a more fruitful task. The author was somewhat surprised to find during the course of the investigation, that many low IQ students displayed a relative problem-solving ability far beyond what one would expect from their overall performance in school subjects and in the standard aptitude tests. It was decided to subject the marked overachievers to a more careful scrutiny with a vew to determining which non-cognitive factors were most in evidence.

1. Selection of Overachievers. The geometry limit scores were converted to standard scores with a mean of 50 and a standard deviation of 10. The IQ scores for the combined group (N=106) were treated in the same way. An overachievement index was defined as follows:

$$A = \frac{\text{Standardized Geometry Score}}{\text{Standardized IQ}}$$

The 20 students showing the highest indices were chosen for special study. The group could be broken down in several ways: (a) 10 boys and 10 girls, (b) 10 grade 10 students and 10 grade 12 students, (c) 12 students with IQ below 106 (the population mean) and 8 students above 106.[1]

2. Sources of Information. Information concerning the students came from many sources. The standardized tests, experimental tests, and questionnaires have already been referred to. Complete school records, dating back to the students' entrance into public school were also used. Each student was interviewed three times in the course of the experimental period. Again, half of this group had been members of the author's mathematics classes for a period of two years. Moreover, the author was aquainted with the home

---

[1]Seven of the 10 students who made continued gains in geometry were included in this group. The three remaining students possessed sufficiently high IQ's scores to lower their overachievement index below the critical value.

[illegible]

[illegible]

[illegible]

[illegible]

life of the majority of these subjects, and had many contacts with the students outside the school situation. When everything is considered, it would appear that there was available, a fairly complete picture of the student, not only as an individual within a mathematics classroom, but also as an adolescent in his daily life.

At this point, two tests, which have not as yet been mentioned will be discussed.

(a) Subjective Utility of Mathematics. The author devised this test to measure the relative subjective value to the student of achievement in mathematics as compared to achievement in verbal subjects. The basic principle employed was the 'indifference curve' (Edwards, 1952), a device used to estimate the rate of exchange employed by the subject to compare the utility of one commodity to the utility of a second commodity. The method is ullustrated in Fig. 45.

The subject was first required to indicate his level of aspiration (i.e., the mark he expected to obtain) in mathematics and in English. The resulting pair of marks were plotted on co-ordinate axes. The student was then asked to imagine that he was in a position to bargain for his mathematics mark. He was asked what loss he would take in his English mark to acquire (1) a gain of one mark in

mathematics, (2) a gain of two marks in mathematics, (3) a gain of 5 marks in mathematics, and so on. Again, he was asked to indicate what increase in the English mark would be necessary to compensate for (1) a loss of one mark in mathematics, (2) a loss of two marks in mathematics, (3) a loss of 5 marks in mathematics.

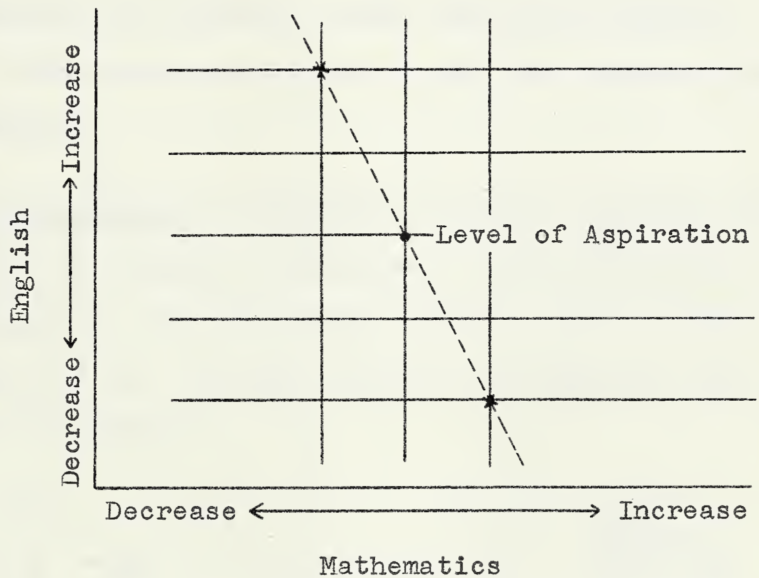


Fig. 45. Utility measurement of ego involvement in mathematics.

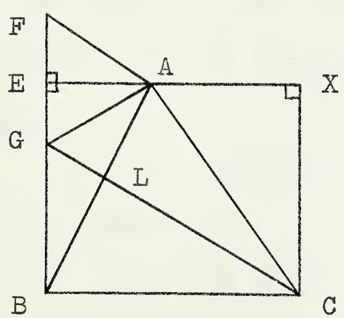
Each answer yields a pair of 'mark' coordinates, and the set of points represents combinations of marks which are of equal 'value' to the student. If the valuation is consistent, a smooth curve may be drawn through the points and the rate of exchange determined.

In the illustration above, the student was willing to

sacrifice two marks in English to gain one in mathematics; again, he required an increase of two marks in English to compensate for a loss of one mark in mathematics. We may infer from this that the mathematics mark has a greater subjective value to the student than the English mark, and we may further infer that the student is ego-involved with mathematics to a greater extent than with English.

This test was applied to all the students in the population.

(b) <u>Perseveration</u>. The author devised a problem intended to measure the student's willingness to persevere at a problem-solving task. The following geometry problem looks plausible enough, but is, in fact, incapable of solution under the stated conditions.



Given: $AC=BC$
$AF=FG=GA$
$\angle AFG=\angle ABC$

Prove: $\angle BFA=\angle AGC$

(note: use A, I, II, III, IV <u>only</u>.)

Fig. 46. Geometry problem employed in the measurement of perseveration.

The problem was assigned on Friday, for the following Monday. It was pointed out to the student that the problem

would provide a real test of problem-solving ability. Each student was required to complete and sign a statement to the effect that he had not received help with the problem and that the solution time recorded was correct.

Only one student in the entire population came to the conclusion that the problem could not be solved. It was of interest, here, to see how long the student would persevere in what had been made out to be an important problem before giving up, or offering an inadequate solution.

3. Comparison Groups. In order to determine non-cognitive correlates which were peculiar to overachievers in geometry, some reference or standard group had to be used. When statements are employed in the following section which indicate that a relatively high incidence of some characteristic was found in the overachiever group, it is to be understood that the incidence was high relative to the population as a whole (N=106).

In a few connections, comparisons were made between the overachievement group and a group of underachievers in geometry, chosen by means of the same index.

The Problem-Solving Syndrome

One cannot have had contact with a group of superior problem solvers in mathematics without having been impressed by their intense motivation. In fact, the necessity of

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

strong motivation can be argued in terms of traditional reinforcement theory. We have previously thought of motivation as a willingness to expend an available amount of energy in a given direction. The expenditure will continue (or response will be made) for a period which is determined by the number of successes (or reinforcements) experienced by the subject.

Comprehension and problem solving differ radically with respect to the frequency of reinforcement per unit of energy expended. Imagine a student following the development of a mathematical argument--the solution of a quadratic equation, let us say. At each step, the individual performs an operation mentally and checks his mental product against the recorded result. Thus there is reinforcement at each step, which takes the form of satisfaction at having followed the argument to the point in question.

In problem solving on the other hand, reinforcement comes only when the problem is completely solved; the intervening mental states, in which alternatives are in competition, provide a feeling of dissatisfaction, which would be opposite in effect to reinforcement.

The overachiever in geometry shows his intense motivation by his willingness to persevere longer than the normally motivated student. In the perseveration test, the 20 overachievers claimed an average of 120 minutes spent in

an attempted solution, compared with the population average of 46 minutes, and the underachievers' average of 18 minutes.

Probing beyond this obviously superior motivation, one naturally searches for its origin. It is well known that the middle class student in the North American culture is strongly motivated toward school achievement, since high school marks pave the way for social and occupational advance. However, this motivation is of the relative variety, since relatively high marks are all that is required. And since respectable marks can be had in geometry, without being able to solve problems, middle class school motivation in itself does not imply the intense ego-involvement seen in the overachiever.

The author suspects that a strong attachment to problem solving in mathematics often originates in a period of the student's life in which he may be suffering from what might be termed social maladjustment. Several pieces of evidence support this point of view.

When the overachievers were contrasted with the whole group, it was found that the former showed a relatively large incidence of high scores[2] on the social relationships

---

[2]A high score on any scale of the Minnesota Counseling Inventory is indicative of poor adjustment in that area.

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

and leadership scales on the Minnesota Counseling Inventory. Moreover, these students seemed to have a very restricted circle of close acquaintances, and were usually to be found by themselves, or in the company of one close friend. They were seldom seen at school dances, were relatively inactive at athletics, and showed little overt interest even at the grade 12 level in members of the opposite sex. There emerges the impression of the geometry overachiever as a somewhat lonely individual who somehow lacks the ability to enter into a sympathetic contact with others.

Since the overachiever tends to feel inadequate in ordinary social intercourse, he shuns school areas in which verbal displays are necessary, and where success depends to some extent on the ability to 'sell oneself' to the group. In mathematics, he finds an area of certainty, in which he can prove his worth merely by producing answers whose correctness is independent of group evaluation.

The student who adopts mathematics as a measure of compensation, establishes a very tenacious hold indeed. He is, of course, intensely ego-involved in the outcome of a particular problem, and for that reason he is willing to persevere. One aspect of his ego-involvement is shown in the fact that mathematics has a relatively high utility value for him.

The overachiever identifies intelligent behaviour

with problem solving, rather than with verbal activities. This may be partly due to an attempt to rationalize his fear of verbal performance in the same way that he rationalizes his avoidance of normal adolescent activities.

In extreme cases, performance in mathematics may become the central prop of the individual's self-concept. Thus a problem which the subject cannot solve becomes a threat. This is evidenced in the intensely motivated student whose pleasure at solving problems is far exceeded by his humiliation and self-disgust when he occasionally fails.

This intense ego-involvement, amounting at times to a logical compulsion, explains the absolute motivation of the overachiever, which far transcends the relative motivation of the ordinary middle class 'achiever'. The overachiever solves problems, not only to earn marks, but because failure to do so threatens the area of competence around which his self-concept is built.

One aspect of absolute motivation--i.e., the desire to be correct rather than to earn marks--is manifested in the overachiever's caution and his refusal to guess. The group under study showed significantly less guessing behaviour on the three available indices of guessing.

Out of this caution, this tendency to sit back and survey the possibilities before leaping ahead, grows the

geometry overachiever's significantly superior scores on all the tests of strategy formation. For strategy formation involves a preliminary assessment of the problem and a bringing into play of that combination of abilities which will yield a solution while minimizing the energy expenditure. Of course, it is because caution and strategy formation are positively related, that Otis IQ, which is largely a measure of speed, itself correlates rather poorly with strategy formation.

In summary, we may say that the overachiever in problem solving often manifests personality traits which indicate that he is attempting to compensate for inadequate social relationships. While the student who possesses these characteristics together with moderately strong innate ability may go on to make satisfactory progress in advanced mathematics, one raises a moot question if he enquires whether, in the long run, the overachiever may not profit more from a better social adjustment, and a poorer problem-solving ability.[3]

---

[3]The 'compensation' explanation of outstanding ability can easily be overstated. The foregoing account has dealt solely with some negative aspects of overachievement, discussed in terms of the rather nebulous concept of 'social maladjustment'. At the same time it must be recalled that we were dealing here with overachievers, who, by definition, could not have been in the extremely high intelligence range. The three remaining students among the 10 who made continued

progress in geometry, were of very high IQ. These students, who in the author's opinion, would have the best chance in succeeding at university mathematics, could not reasonably be classified as cases of 'compensation'. It would certainly seem to be plausible to argue that over and above the negative aspects of the mediocre-IQ, high-geometry over-achiever, emerged the somewhat intangible aspects of the ability of the high-IQ geometry achiever aspects which merge at the higher levels with the creativity of the professional mathematician.

CHAPTER XVI

SUMMARY OF THEORY AND FINDINGS

Theory

This thesis presented an energy theory of mental operations, and applied it to the problem of mathematical ability.

When thinking is considered from a phylogenetic and maturational point of view, it seems apparent that the brain has evolved to serve the principal function of facilitating homeostatic equilibrium. This function is accomplished by extending the region within which the organism can anticipate possible disruption of its equilibrium. In short, the organism is able to plan for, and deal with future events because it acquires the power to construct and manipulate systems imaginatively.

When one attempts to explain how the brain acquires its anticipatory powers, one immediately encounters energy considerations. While the brain is electrically active from birth onward, its early activity is disorganized, and the extent of 'central' as opposed to 'environmental' control of cerebral processes is vastly limited. During this time the brain is under constant bombardment from external stimuli, and there occurs a mapping of the external world on its association layers. Gradually, objects in the exterior

environment come to be represented internally by a certain complexity of neural organization and activity.

At some point in the maturation of the brain, the representation of the exterior object can be centrally activated without assistance from the exterior energy source. Similarly, representatives can be mentally manipulated, i.e., manipulated in 'imagination'. Since this imaginative activity is equivalent to the activity which would be caused by an external source of energy, we are obliged to postulate as well, an internal source of energy. At the same time, we are obliged to move from a neural to a psychic level of description, and to invoke the concepts of ego and cathexis. While one cannot equate cathectic and physiological energies, the obvious interaction between neural and psychic events renders safe the assumption that some part of the two energies are isomorphic under operations of increase and decrease.

The amount of residual hypercathexis, and hence, the possibility of central control, depends upon the contemporary state of activity of the organism's configuration of needs. In this respect, the hypercathexis level seems to bear an inverse relationship to the strength of needs.

When the interaction of organisms is considered, a third, or social level of description is reached from which arises the notion of self-concept. Since the enhancement of

the self-concept operates as a need, it is capable of causing a redistribution of the cathexis-hypercathexis balance.

Two primary schema were proposed for centrally-controlled thought. The first, which might be called the 'undivided energy' schema (Type A), consisted of a sequence of operation-storage couplets. The second, or 'divided energy' schema (Type B) occurred in situations where logical or utility choices were to be made.

The axiomatic system in mathematics was next considered, and levels of ability were distinguished in terms of the student's task within the system. Comprehension of mathematics, denoted Level I ability, seemed to fit the 'undivided energy' schema; in fact, one is tempted to argue that mathematics has evolved until it corresponded to this basic mode of mental operation. The energy theory offered no reason for supposing that comprehension should possess natural upper bounds. However, it was seen that low motivation--understood as a utility partition of available energy--could result in an application of insufficient hypercathexis to the comprehension task, and that upper bounds would appear in this case. Again, it was postulated that adverse interaction (at the third level of description) could pose a threat to the self-concept, and that this in turn, could result in a redistribution of the cathexis-hypercathexis balance, resulting in the appearance of

induced upper bounds.

A mathematical problem was said to originate when a gap occurred in the deductive sequence, and when the subject developed, or was presented with, a systematic method of filling the gap which generated successive choice points. Thus problem solving (Level II ability) corresponds to the 'divided energy' schema, and the theory would predict natural upper bounds. Again, low motivation and group pressure were seen to be capable of lowering the potential upper bounds.

The third, or creative function (within the axiomatic system) had to do with the creation of axiom systems and the extension of existing systems in fruitful directions. Since our knowledge of this area is inadequate to form the basis of a science, no attempt was made to extend the theory to it.

Experimental Findings

The experiments were designed to corroborate specific aspects of the theory. Consequently, two major areas of experimentation emerged, comprehension (Level I) and problem solving (Level II).

1. Studies at Level One. Two tests of comprehension were designed for experimental use. One test employed number series constructed by the theory of finite differences. The second test employed permutations on four elements. The

principal results may be summarized as follows:

(a) In both cases, the operation$\longrightarrow$storage$\longrightarrow$operation-on-storage sequence possessed a time function which was noticeably exponential, for the first few operation-storage couplets at least. There was some evidence in the case of number series that the curve overestimated the required comprehension time beyond the fifth couplet.

(b) While there was no evidence to show the necessity of natural upper bounds in comprehension, artificial upper bounds were induced in a group-testing situation which was carefully controlled to exert pressure on the individual. When the time limits were controlled in such a manner that a gradient of decreasing group pressure was created, guessing behaviour declined significantly. Moreover, when group pressure was completely removed, the subject made significant gains in their upper bounds.

(c) The two measures of the speed of performing simple mental operations correlated highly with Otis IQ and with each other. It would seem plausible to infer a relationship between energy level and speed of operations.

(d) No significant differences were observed in the rate of performance of mental operations between grades 10 and 12 groups matched on Otis IQ. It would seem, then, that once the stage of formal mental operations is reached, further maturational processes do not increase the maximum

energy level.

(e) A comparison of the grade 10 and grade 12 matched groups showed the latter group to be significantly superior on two tests of strategy formation. At the same time, while the relationship between strategy formation and Otis IQ seemed to be a function of task difficulty, this relationship appeared to be generally low for both tests.

(f) A study of the comprehension of topological concepts revealed that in a situation employing standard mathematical material, comprehension time again appeared to be on closer to an exponential, than to a linear, function of deductive length. It also appeared that grade 12 students were capable of comprehending topological concepts, provided that adequate representatives were available, and sufficient time was allowed for storage.

2. Studies at Level Two. The major emphasis was given to the study of geometry problems, or problem solving under finite uncertainty. The principal findings are listed below.

(a) The deductive and uncertainty measures of problem difficulty both gave fairly accurate estimates of the experienced problem difficulty. The uncertainty model in particular gave considerable insight into the difficulty encountered by the student in his attempts to solve problems.

(b) When the deduction and uncertainty measures were combined to form a single index of problem difficulty, the division of problems on the basis of this index yielded categories which could be given conventional descriptive verbal labels.

(c) In terms of the difficulty categories employed, the performance of most geometry students seemed not only to remain stable relative to the group over a long period of time, but to be predictable by the end of a three week 'trial run' in geometry. In particular, the student who was only able to solve direct applications of the propositions by the end of the first three weeks, seemed from that time onward, only to add propositions to his repertoire without increasing the depth to which he could analyse problems.

(d) Several pieces of evidence supported the 'boundedness' theory of problem solving. All but 5 of the grade 10 students showed predictable upper bounds in performance within the 11 week period. Again, repetition of problems did not seem to cause a rise in the incidence of correct solutions. Finally, a comparison of the matched grade 10 and grade 12 groups showed that once the grade 10 group had reached their upper bounds, no significant differences could be found in the frequency of solutions of common problems.

At the same time, a small minority of students

(roughly 10%) seemed to make continued gains in raising their upper bounds. These students employed a refinement of the standard strategy in which they made a further classification of decision types. Each type was associated with a particular geometrical configuration. The refined strategy caused a decrease in the problem uncertainty which in turn resulted in a decrease in subjective problem difficulty. Consequently, a fixed amount of energy would establish a higher upper bound with the refined strategy than with the fixed strategy. The student who made continued progress in problem solving seemed to effect a continuous refinement of strategy.

(e) An analysis of behaviour near and above the upper bound showed that while the individual was able to produce a complete analysis below his upper bound, the analysis gradually broke down when he attempted problems above his upper bound, and the individual then made assumptions which had the effect of decreasing the problem difficulty. The subject gradually moved from an attempt to be logical, to an attempt to be plausible, and finally offered an outright guess in an apparent attempt to bluff his way through. While some students offered few solutions when they found that the problems had become too difficult for them, yet there was enough guessing in evidence that a majority of students gave more incorrect than correct solutions.

(f) In problem solving under infinite uncertainty, little evidence of strategy formation was found within the experimental period. Rather, the student brought predetermined strategies into the testing period. These fixed strategies, together with some momentary expectations generated by the structure of the battery, determined the solutions offered by the student. Here again, an alarming number of students blindly applied inadequate strategies, and by ignoring part of the data, obtained solutions which were clearly contradicted by the complete facts.

(g) Four measures of strategy formation showed positive and statistically significant correlations at the grade 12 level. Moreover, strategy formation was not highly related to Otis IQ. In particular, geometry problem solving scores and Otis IQ correlated to the extent of .4 to .5.

(h) A multiple coefficient of correlation, using geometry as the criterion, and (1) Otis IQ, (2) D.A.T. Numerical, (3) D.A.T. Spatial, (4) Watson-Glaser, as predictors, amounted to .63 for grade 10 and .82 for grade 12. In both cases, the D.A.T. Spatial was not only the best individual predictor, but also showed the highest weight, while the Otis ranked third and fourth (in weight) for grades 10 and 12 respectively. It was inferred from the grade 10 predictors that an accurate prediction of geometry performance could best be made by the preliminary trial

method.

(i) An investigation of non-cognitive correlates of problem solving ability was carried out for the whole population. Sex differences in the sample were observed only to the extent that grade 12 girls showed a sensitivity to the fact that mathematics is, by social definition, a 'man's' area.

It was apparent that the successful problem solver developed early expectations of success, partly through his public school experiences in mathematics, and partly from parental experiences in, and attitudes toward, mathematics.

The study of the overachiever in geometry revealed that his most outstanding characteristics were an intense motivation which manifests itself in a willingness to persevere, a cautious attitude toward guessing, and a strong ego-involvement with the outcome of problem solving. There was some evidence to support the hypothesis that the problem solving drive may originate in a period of social maladjustment.

Implications for the School

It seems to the author that a simplified energy theory of mental operations could be of considerable use to the teacher of high school mathematics. Of particular importance, is the notion of the non-linear time function of the operation-storage couplet, and the consequent necessity

of adequate storage time. When the teacher accepts the idea that his teaching of the comprehension of mathematics ought to be primarily concerned with the demonstration of operations, then he will have a theoretical basis from which to work--something which is, at present, lacking. It will mean that the verbalization of operational rules will be reduced to a minimum, while physical demonstrations will be maximized.

The teaching of the operational rules of signed numbers is a case in point. Current textbooks offer verbal rules for the subtraction of negative numbers which are completely meaningless from the point of view of operations performable on physical objects. A method more in keeping with the physical basis of comprehension--and one which even the slowest student seems to be able to comprehend--explains subtraction in terms of performable movements on a line diagram.

At the same time, the importance of group influence on comprehension cannot be ignored. In a class with wide differences in mathematical comprehension rates, the slow student is not only expected to produce answers, but is made constantly aware of his slow performance. Under these circumstances, we must expect to see a great deal of that illogical behaviour which can be subsumed under the general heading of 'guessing'.

In any case, a theory of logical behaviour, which

treats the individual as though he were thinking in a social vacuum, is thoroughly inadequate as an explanation of observed behaviour. Perhaps the whole theory of the 'logical' man is unrealistically idealistic. In the vast part of the individual's life, the desire or need, to be logical is subordinate to more fundamental drives. If we should find it desirable to attempt to maximize the student's logical output, then careful controls must be established within the classroom to allay these more fundamental needs.

The present study also has something to say in connection with problem solving. The geometry course of the high school has often been attacked in the past, chiefly on the grounds that no observable transfer has been manifested with regard to logical behaviour. In other words, the contention has been that geometry is more or less ineffective as a general method of developing the ability to reason correctly.

The author would contend that geometry may, in fact, be harmful to a majority of the students who are at present obliged to study it. It is difficult to see how anything of positive value can emerge from a training in which the student, through assumptions, circular reasoning, and plain bluffing, reduces complex problems to a level at which he can deal with them. The present study was conducted under conditions of organization, motivation, and individual

attention to students' work which can scarcely be equalled in ordinary day-by-day teaching. The author suspects that logical output in the typical classroom under a not-too-well-qualified teacher, may be very small indeed. Unfortunately, the classroom teacher is in no position to assess the daily products of each student's thinking.

It we continue to teach geometry in the high school, some kind of strict ability division must be effected within the classroom, so that each student attempts problems which he has at least a fair chance of solving. This thesis shows how such a division is possible, provided that the teacher is allowed sufficient time to make the necessary preliminary assessments.

Again, more emphasis must be given to the notion that many real-life prlblems, and perhaps the vast majority of important ones, do not readily lend themselves to complete and final analysis. In light of the fact, the most important mental characteristic which we should attempt to inculcate in the student is the desire to suspend judgment in the face of problems which his mental powers are inadequate to deal with. This may well involve a reshifting of emphasis in the assignment of school marks, so that the individual's performance is not judged solely on the number of solutions or part solutions that he produces.

Finally, if the basis of the tenacious attachment to problem solving lies--as the author suspects--in a type of social maladjustment, the educator must decide whether the extremely ego-involved problem solver might not ultimately profit more from an extension of his social adequacy.

Suggestions for Further Research

(1) The effect of group influences on individual thought and performance remains a practically untouched area. It is scarcely necessary to elaborate its implications for the classroom. It may turn out that the notion that problem solving can be taught in a group situation is psychologically invalid.

(2) In view of the 'non-linearity of comprehension' hypothesis proposed in this thesis, the author suggests that a series of experiments be carried out at the university level. The hypothesis is put forward here, that university mathematicians assume and lecture on the linearity hypothesis, and that comprehension of the typical mathematics lecture will be found to be very small.

(3) It would seem worthwhile to extend the concept of strategy. The energy theory postulates the existence of a primitive source of mental energy, which is reflected in the speed of performing operations at a low order of difficulty. Strategy formation comes into play when the

individual attacks problems which tend to recur in his experience. It involves the development of an approach which brings the particular combination of the individual's abilities into play which will minimize the energy expenditure necessary to obtain a solution.

While strategy formation would seem to be an important aspect of intelligent' behaviour, there seems to be little in the standard IQ test which assesses this ability.

(4) The upper bound approach to mathematical performance would seem to be a useful one, and capable of extension at the high school and university levels. This point of view is somewhat at variance with current research in educational psychology in which emphasis is put on discovery of methods of improving on what the individual <u>can</u> do rather than finding out what levels are closed to him because of limitations in innate ability.

The author believes that every student is capable of <u>comprehending</u> some part of mathematics. However, differences in individual rates of comprehension are so great that the slower student cannot possibly follow the development in the time allotted, since school teachers operate under the requirements of a rigid syllabus and must adjust their exposition so that a fixed quantity is covered in a given time.

Under these circumstances, there are many students who

are incapable of profiting from the study of mathematics under normal classroom time requirements--even when the best methods of instruction that are currently known are used. It would seem feasible, and desirable, to be able to estimate that in a specified class situation, a given individual will not progress beyond a determinable level in mathematics. This thesis has shown how such estimation can be made in geometry, and a similar technique would undoubtedly be applicable to other mathematical areas.

The next few years may witness great changes in the mathematics curriculum at the high school level. There is some danger in assuming that a change in subject matter will, in itself, remedy the ills of mathematical education. The real problem does not lie here, but in discovering the psychological processes which underlie mathematical ability. It is to these underlying processes that our research must be directed.

GLOSSARY

<u>attention-cathexis</u>: the application of hypercathexis in a given direction.

<u>cathexis</u>: the primitive energy which powers the organism's effort to satisfy its basis needs. The term was originally used by Freud who equated it to physiological energy. More recently, a critical evaluation of the term has resulted in a more cautious interpretation of its status as energy.

<u>cell assembly</u>: the hypothesized elementary electrical circuit which are formed in the brain as a result of a continued pattern of stimulation from the environment.

<u>central facilitation</u>: the process of activating or intensifying the activity of neural structures by an energy transformation which does not emanate directly from an external source.

<u>drive</u>: an aroused condition of the organism based upon deprivation or noxious stimulation, including tissue needs, drug or hormonal conditions, and specified internal or external stimuli.

<u>ego</u>: In Freud's tripartite division of the personality, the part corresponding most nearly to the perceived self, the controlling self which holds back the impulsiveness of the id in the effort to delay gratification until it can be found in socially approved ways.

<u>homeostasis</u>: an optimal level of organic function, maintained by regulatory mechanisms known as "homeostatic mechanisms", e.g., the mechanisms maintaining a uniform body temperature.

<u>hypercathexis</u>: the cathexis which remains over and above the amount required for the primary process. An accumulation of hypercathexis results from the increased efficiency of the organism which accompanies its maturational growth and experiences in the environment. Hypercathexis is employed by the ego in the performance of the reality testing function and is often used synonymously with the term 'psychic energy'.

incentive: a tangible goal object which provides the stimuli that lead to goal activity.

Level I ability: the comprehensional aspect of mathematical performance relative to an axiomatic system; the performance of the permissible operations on the primitive elements of the system and the storage of successive reference points.

Level II ability: the gap-filling aspect of mathematical performance relative to an axiomatic system. Level II ability (problem solving) implies the existence of a gap-filling strategy which generates successive choice points.

Level III ability: the creative aspect of mathematical performance relative to an axiomatic system. It implies the creation of new axioms or axiom systems and the extension or inter-relating of established systems.

level of aspiration: a goal that the individual sets as something he expects to achieve or strives to achieve. Reaching that goal is interpreted by him as success, falling short as failure.

phase sequence: a network of cell assemblies of varying degrees of complexity.

primary reference: the source of the subjective feeling of the 'correctness' or 'truth' of elementary mathematical operations which follows from the performance of internalized physical operations upon the primitive representatives of a system.

secondary reference: the source of the subjective feeling of the correctness of a mathematical development which proceeds by the application of rules whose truth has previously been established by primary reference.

set: a preparatory adjustment or readiness for a particular kind of action or experience. Used as a synonym for Einstellung.

storage: temporary storage refers to the maintenance of the state of activity in a neural structure above a critical operational level. Permanent storage refers to the modification of the neural structure (synaptic growth) to the extent that central activation of the neural structure becomes possible.

storage time: the interval between the initial activation of a phase sequence system and its employment in further mental operations.

strategy: the term implies three conditions: (a) the repetition of events of certain types, (b) an awareness on the part of the individual of the possible range of events, (c) a more-or-less clearly formulated plan on the part of the individual to deal with the occurrence of a particular event in a specific way.

subjective probability: the relative frequency with which an individual expects an event to occur, in contrast to objective frequency, the relative frequency with which it actually does occur.

Type A schema: a form of mental operation which in its purest form is represented by the operation-storage-operation-on-storage schema. It may be said to be an undivided energy model in that no competing mental states are encountered.

Type B schema: a form of mental operation in which attention-cathexis is divided between competing alternatives of the 'utility' or 'logical' type.

upper bound: a sequence of numbers is said to have an upper bound if it is possible to specify some number which is not exceeded by any member of the sequence.

BIBLIOGRAPHY

BIBLIOGRAPHY

Allandoerfer, C. B. Deductive methods in mathematics; in Insights into modern mathematics. Twenty-third yearbook of the National Council of Teachers in Mathematics. Washington, D.C., National Council of Teachers of Mathematics, 1957, 65-100.

Anderson, G. L. Quantitative thinking as developed under connectionist and field theories of learning. In Learning theory in school situations. University of Minnesota Studies in Education, No. 2. Minneapolis, Univ. of Minn. Press, 1949, 40-73.

Anderson, T. W. Probability models for analysing time changes in attitudes. In P. F. Lazarsfeld, (ed.), Mathematical thinking in the social sciences. Glencoe, Ill., Free Press, 1954.

Ashby, W. R. Design for a brain. New York, John Wiley & Sons Inc., 1952.

Ashby, W. R. An introduction to cybernetics. New York, Wiley, 1956.

Barakat, M. K. A factorial study of math. abilities. Brit. J. Psychol., Stat. Sec., 4, 1951, 137-156.

Bartlett, Sir Frederick. Thinking, a psychological and social study. London, Routledge and Paul, 1957.

Barr, F. Urban and rural differences in ability and attainment. Educ. Res. (Brit.), 1, 1959, 49-59.

Berlyne, D. E. Recent developments in Piaget's work. Brit. J. Psychol., 48, 1957, 1-12.

Bexton, W. H., Heron, W., Scott, T. H. Effects of decreased variation in the sensory environment. Canadian J. Psychol., 8, 1954, 70-76.

Biggs, J. B. The development of number concepts in young children. Educ. Res. (Brit.), 1, 1959, 17-34.

Binet, A. The psychology of reasoning, Chicago, Open Court Publishing Co., 1907.

Blackwell, A. N. A comprehensive investigation into the factors involved in mathematical ability in boys and girls. Brit. J. educ. Psychol., 10, 1940, 143-153, 212-222.

Boole, G. An investigation of the laws of thought. New York, Dover Publications, 1951.

Bowers, J. E. The difficulty of one-step arithmetical problems in relation to the type of fundamental number operations involved. Unpublished Ed.D. thesis, University of Toronto, 1955.

Braithwaite, R. B. Scientific explanation, a study of theory, probability and law in science. Cambridge, University Press, 1953.

Breuer, D. The problem of quantification in psychological sciences. Psychol. Rev., 65, 1958, 8-13.

Brown, K. E. Research in teaching high school mathematics. Math. Tchr., 51, 1958, 578-585.

Brownell, W. A. Psychological consideration in the learning and teaching of arithmetic. In The teaching of arithmetic, Tenth Yearbook of the National Council of Teachers of Mathematics. New York, Teachers College, Columbia University Bureau of Publications, 1935, 1-32.

Brownell, W. A., et al. Arithmetic in grades one and two: A critical summary of new and previously reported research. Duke University Research Studies in Education, No. 6, Durham, N.C., Duke University Press, 1941.

Brownell, W. A. A problem solving process that succeeds arithmetic we need series, Boston, Ginn & Co., 1956.

Bruner, J. S., Goodnow, J. J., Austin, G. A. A study of thinking. New York, Wiley, 1956.

Bruner, J. S. et al. Contemporary approaches to cognition. Cambridge, Harvard University Press, 1957.

Buckingham, B. R. & MacLatchy, Josephine H. Number abilities of children when they enter grade one. Report of the society's committee on arithmetic. Twenty-ninth Yearbook of the National Society for the Study of Education. Bloomington, Ill., Public School Publishing Co., 1930, 473-524.

Buswell, G. T. The psychology of learning in relation to the teaching of arithmetic. In The teaching of arithmetic, Fiftieth Yearbook of the National Society for the Study of Education, Part II. Chicago, National Society for the Study of Education, 1951, 143-154.

Buswell, G. T. Needed research in arithmetic. In The teaching of arithmetic, Fiftieth Yearbook of the National Society for the Study of Education, 1951, 282-297.

Buswell, G. T., with the cooperation of Kersh, B. Y. Patterns of thinking in solving problems. University of California Publications in Education, 12, 11, Berkeley and Los Angeles, University of California Press, 1956, 63-148.

Buswell, G. T. Educational theory and the psychology of learning. J. educ. Psychol., 1956.

Carnap, R. Introduction to semantics. Cambridge, Mass., Harvard University Press, 1942.

Cattell, R. B. Personality traits associated with abilities: With verbal and math. abilities. J. educ. Psychol., 36, 1945, 475-486.

Cattell, R. B. Factor analysis. New York, Harper, 1952.

Cherry, E. C. On the validity of applying communication theory to experimental psychology. Brit. J. Psychol., 48, 1957, 176-188.

Clarke, S. C. T. Clear thinking in geometry. Unpublished M.Ed. thesis, University of Alberta, 1943.

Cohen, J. C., Dearnaley, E. J. and Hansel, C. E. M. Measures of subjective probability. Brit. J. Psychol., 48, 1957, 271-275.

Colby, K. M. Energy and structure in psychoanalysis. New York, Ronald Press Co., 1955.

Coleman, R. H. An analysis of certain components of mathematical ability and an attempt to predict mathematical achievement in a specific situation. Unpublished Ph.D. thesis, Indiana, 1956, Reported in Dissertation Abstracts, 1, 1951, 2-62.

Crawford, D. H. An introductory study of the consistency of marking mathematics papers. Alberta J. Educ. Res., 5, 1959, 31-40.

Cronbach, L. J. A generalized psychometric theory based on information measure. Urbana, Ill., College of Education, 1952.

Cronbach, L. J. and Gleser, C. G. Psychological tests and personal decisions. Urbana, Ill., University of Illinois Press, 1957.

Croxton, F. E., and Cowden, D. J. Applied general statistics. Englewood Cliffs, N. J., Prentice-Hall, 1956.

Davis, K. Human Society. New York, Macmillan Co., 1949.

Davis, R. C. The domain of homeostasis. Psychol. Rev. 65, 1958, 8-13.

Dewey, J. How we think. Boston, Heath, 1920.

Dodes, I. A. The science of teaching mathematics. Math. Tchr., 46, 1953, 157-166.

Douglas, C. K. Relation of high school to college mathematics marks. Sch. Rev., 44, 1936, 615-619.

Duncker, K. On problem solving. Psychol. Monogr., 58, No. 5, 1945.

Dyer, H. S., Kalin, R., Lord, F. M. Problems in mathematical education. Princeton, N.J., Educational Testing Service, 1956.

Edwards, W. The theory of decision making. Psychol. Bull. 51, 1954, 380-417.

Fehr, H. F., McMeen, G., and Sobel, M. Using hand-operated computing machines in learning arithmetic. Arith. Tchr. 3, 1956, 145-50.

Fehr, H. F. Theories of learning related to the field of mathematics. Twenty-first Yearbook of the National Council of Teachers of Mathematics, Washington, D.C., National Council of Teachers of Mathematics, 1953, 1-41.

Fehr, H. F. Breakthrough in mathematical thought, Math. Tchr. 52, 1959, 15-20.

Fleming, C. M. Class size as a variable in the teaching situation. Educ. Res. (Brit.), 1, 1959, 35-47.

Freeman, G. L. The energetics of human behaviour. Ithaca, N.Y. Cornell University Press, 1948.

Freud, S. The basic writings of ... New York, Modern Library, Inc., 1938.

Freud, S. An outline of psychoanalysis. New York, W. W. Norton & Co. Inc., 1949.

Gagne, W. Problem solving and thinking. Annual Rev. of Psychol., 10, 1959, 147-171.

Garrett, H. E. Statistics in psychology and education. New York, Longmans, Green & Co., 1953.

Gibb, E. G. A review of a decade of experimental studies which compared methods of teaching arithmetic. J. educ. Res. 46, 1953, 603-708.

Gilchrist, J. C. Social psychology and group processes. Annual Rev. of Psychol., 10, 1959, 233-263.

Girshick, M. A. An elementary survey of statistical decision theory. Rev. educ. Res., 5, 1954, 448-467.

Glaser, R., Schwartz, P. A. Scoring problem solving items by measuring information. Educ. Psychol. Measmt., 14, 1954, 665-670.

Glaser, R., Damrin, D. E. and Gardiner, F. The tab item. Educ. Psychol. Measmt. 14, 1954, 283-294.

Glennon, V. J. and Hunnicut, C. W. What does research say about arithmetic? Washington, D. C., Association for Supervision and Curriculum Development, 1952, 45.

Guilford, J. P. Factor analytic study of navy reasoning tests. Educ. Psychol. Measmt. 14, 1954, 301-325.

Guilford, J. P. The structure of the intellect, Psychol. Bull., 53, 1956, 267-293.

Gunderson, Agnes G. Number concepts held by seven-year-olds. Math. Tchr., 33, 1940, 18-24.

Hadamard, J. The psychology of invention in the mathematical field. Princeton, N.J., Princeton Univ. Press, 1945.

Haggard, E. A. Intraclass correlation with the analysis of variance. New York, Dryden Press, 1958.

Hall, C. S., and Lindsey, G. Theories of personality. New York, Wiley, 1957.

Hamley, H. R. Relational and functional thinking in mathematics. Ninth Yearbook of the National Council of Teachers of Mathematics, Teachers College, Columbia University, New York, 1934.

Hartree, D. R. Numerical analysis. Oxford, The Clarendon Press, 1952.

Hartung, M. L. Motivation for education in mathematics. Twenty-first Yearbook of the National Council of Teachers of Mathematics, Washington, D.C., National Council of Teachers of Mathematics, 1953, 42-68.

Havighurst, R. J. Can mathematicians and educationists cooperate? Math. Tchr. 30, 1937, 211-213.

Hebb, D. O. The organization of behaviour. New York, John Wiley Inc., 1949.

Henderson, K. B. and Pingry, R. E. Problem solving in mathematics. Math. Tchr., 43, 1950, 228-269.

Hertzberg, F. and Lapkin, M. A. A study of sex differences on the primary mental abilities test. Educ. Psychol. Measmt., 14, 1954, 687-689.

Hick, W. E. Information theory and intelligence tests. Brit. J. Psychol., 42, 1951, 183-194.

Hilbert, D. and Ackermann, W. Principles of mathematical logic. New York, Chelsea Publishing Co., 1950.

Hilgard, E. R. Theories of learning. New York, Appleton-Century-Crofts, 1956.

Hilgard, E. R. Introduction to psychology. New York, Harcourt-Brace, & Co., 1957.

Hills, J. R. Relationship between certain factor analysed abilities and success in college mathematics. Educ. Psychol. Measmt., 17, 1957, 615-622.

Hoel, L. An experiment in clinical procedures for arithmetic. In Emerging practices in mathematics education. Twenty-second Yearbook of the National Council of Teachers of Mathematics, Washington, D.C., National Council of Teachers of Mathematics, 1954, 222-232.

Hoffman, L. R. Homogeneity of member personality and its effect on group problem solving. J. abnorm. soc. Psychol. 1958, 27-33.

Humphrey, C. The problem of the direction of thought. Brit. J. Psychol., 30, 1940, 183-196.

Hunter, I. A. The solving of three-term series problems. Brit. J. Psychol., 48, 1957, 286-299.

Huntington, E. V. The fundamental propositions of algebra, in J. W. A. Young (ed.), Monographs on topics of modern mathematics, New York, Dover Publications, 1955.

Inhelder, B. and Piaget, J. The growth of logical thinking from childhood to adolescence. London, Rutledge and Regan Paul, 1958.

Jacobson, J. E. The electro-physiology of mental activities. Amer. J. Psychol., 44, 1932, 667-694.

Johnson, D. M. Problem solving and symbolic processes, Annual Rev. of Psychol., 1, 1950, 297-309.

Kattsoff, L. O. A philosophy of mathematics. Aimes, Iowa, Iowa State College Press, 1949.

Kelland, N. The prestige value of certain aspects of the educational program in Alberta composite high schools, M.Ed. thesis (in preparation), University of Alberta, 1959.

Kenner, M. R. Meta-mathematics and the modern conception of mathematics. Math. Tchr., 51, 5, 1958, 350-357.

Kenney, J. F. and Keeping, E. S. Mathematics of statistics, Part 2. New York, Van Nostrand, 1954.

Khinchin, A. I. The mathematical basis of information theory. New York, Dover Publications, 1957.

Kinney, L. B. and Purdy, C. R. Teaching mathematics in the secondary school. New York, Rinehart & Co. Inc., 1952.

Kleene, S. C. Introduction to meta-mathematics. New York, Van Nostrant Co., Inc., 1952.

Knight, F. B. Learning theory in arithmetic. Report of Committee of National Society for the Study of Education Part One, Twenty-ninth Yearbook of the National Society for the Study of Education, Chicago, University of Chicago Press, 1930, 145-267.

Koehler, W. Gestalt psychology. New York, Liveright, 1929.

Lazerte, M. E. The development of problem solving ability in arithmetic. Toronto, Clarke, Irwin & Co., Ltd., 1933.

Lazarsfeld, P. F. Mathematical thinking in the social sciences, Glencoe, Ill., Author, 1954.

Lee, D. M. A study of specific ability and attainment in mathematics. Brit. J. educ. Psychol., 25, 1955, 178.

Leeper, R. Lewin's topological and vector psychology, a digest and a critique. Eugene, Oregon, University of Oregon Press, 1943.

Lewin, K. Principles of topological psychology. New York, McGraw-Hill, 1936.

Lewis, H. An experiment in developing critical thinking through the teaching of plane demonstrative geometry. Math. Tchr., 43, 1950, 411-413.

Lord, F. E. A study of the spatial orientation of children. J. educ. Res., 34, 1941, 481-505.

Lorimer, F. The growth of reason. New York, Harcourt Brace & Co., 1929.

MacKay, J. G. An analysis of the abilities of difficulties in geometry 1 as taught in Alberta high schools. Unpublished M.A. thesis, University of Alberta, 1943.

McConnell, T. R. Recent trends in learning theory and their application to the psychology of arithmetic. In arithmetic in general education, Sixteenth Yearbook of the National Council of Teachers of Mathematics. New York, Teachers College, Columbia University, Bureau of Publications, 1944, 143-154.

McNemar, Q. Psychological statistics. New York, John Wiley & Sons Inc., 1955.

McKellar, D. Imagination and thinking, a psychological analysis. London, Cohen and West, 1957.

Marshack, J. M. Probability in the social sciences, in P. F. Lazarsfeld, Mathematical thinking in the social sciences. Glencoe, Ill., Free Press, 1954.

Massimiano, C. C. The influence of the study of plane geometry on critical thinking. Math. Tchr. 49, 1956, 151-153.

Max, L. W. An experimental study of the motor theory of consciousness III action-current responses in deaf-mutes during sleep. Sensory stimulation and dreams. J. Comparative Psychol., 193, 1935, 469-486.

Mead, G. H. Mind, self, and society. Chicago, University of Chicago Press, 1934.

Meder, A. E. (JR.) The maneuvers in set thinking. Math. Tchr., 51, 1958, 578-585.

Michael, R. E. The relative effectiveness of two methods of teaching certain topics in ninth grade algebra. Math. Tchr. 42, 1949, 83-87.

Michael, W. B., Zimmerman, W. S., and Guilford, J. P. An investigation of the nature of spatial relations and visualization factors in two high school samples. Educ. Psychol. Measmt. 11, 1951, 511-577.

Michael, W. B. Differential testing of high level personnel. Educ. Psychol. Measmt. 17, 1957, 475-489.

Miller, G. A. and Frick, F. C. Statistical behaviouristics and sequences of responses. Psychol. Rev. 56, 1949, 311-324.

Miller, G. A. and Selfridge, J. A. Verbal context and the recall of meaningful material. Amer. J. Psychol. 63, 1950, 173-185.

Miller, K. E. Einstellung rigidity, intelligence and teaching methods. Brit. J. Psychol. 48, 1957, 32-48.

Milner, P. M. The cell assembly, Mark II, Psychol. Rev., 64, 1957, 242-253.

Oldham, H. W. A psychological study of mathematical ability. Brit. J. educ. Psychol., 7, 1937, 269-286 and 8, 1938, 16-28.

Pallensen, L. C. An investigation of the reasoning processes employed in a first course in demonstrative geometry. Unpublished M.Ed. thesis, University of Alberta, 1943.

Patrick, C. Creative thought in artists, J. Psychol., 4, 1937, 35-73.

Petrie, P. A. et al. Intermediate mathematics, book four. Toronto, Copp clark, 1954.

Phillips, H. B. Calculus. New York, John Wiley & Sons, Inc., 1927.

Piaget, J. The psychology of intelligence. New York, Harcourt, Brace, 1950.

Piaget, J. The biological problem of intelligence, In Rapaport, The organization and pathology of thought. New York, Columbia University Press, 1951, 174-192.

Piaget, J. Principal factors determining intellectual evolution from childhood to adult life. In D. Rapaport, The organization and pathology of thought. New York, Columbia University Press, 1951, 154-174.

Piaget, J. The child's conception of number. London, Rutledge & Paul, 1952.

Plank, N. Observations on attitudes of young children toward mathematics. Math. Tchr. 43, 1950, 252-263.

Polya, G. Mathematics as a subject for learning plausible reasoning. Math. Tchr. 52, 1959, 7-10.

Price, H. V. Needed research in the teaching of science and mathematics. Sch. Science and Math. 57, 1957, 389-390.

Quastler, H. Information theory in psychology, Glencoe, Ill., Author, Free Press Publishers, 1955.

Rapaport, D. The organization and pathology of thought. New York, Columbia University Press, 1951.

Rignano, E. The psychology of reasoning, New York, Harcourt, Brace & Co., 1923.

Roberts, C. A. Sex differences in children, Report series #2, Ottawa, Mental Health Division, 1956.

Robinson, G. de B. The foundations of geometry. Toronto, University of Toronto Press, 1946.

Rogers, A. L. Experimental tests of mathematical ability and their prognostic value. New York, Columbia University Press, 1918.

Rosenbaum, R. A. Sweet are the uses of generality. Math. Tchr., 51, 1958, 87-95.

Rourke, R. E. K. Some implications of twentieth century mathematics for high schools. Math. Tchr. 51, 1958, 78-87.

Russel, D. H. Children's thinking. Boston, Ginn & Co., 1956.

Russell, N. M. Arithmetical concepts of children. J. educ. Res., 30, 1936, 647-663.

Savage, L. J. Theory of statistical decision. J. Amer. Statis. Assoc., 46, 1951, 55-67.

Savage, M. E. Knowledge of educational research possessed by Ontario high school mathematics teachers. Unpublished M.Ed. thesis, University of Toronto, 1957.

Schunert, J. The association of mathematical achievement with certain factors resident in the teacher, in the teaching, in the pupil, and in the school. J. exp. Educ., 19, 1951, 219-238.

Shannon, C. E. and Weaver, W. The mathematical theory of communication. Urbana, Ill., University of Illinois Press, 1949.

Siegel, S. Nonparametric statistics for the behavioural sciences. Toronto, McGraw-Hill, 1956.

Siegel, S. Level of aspiration and decision making. Psychol. Rev., 64, 1957, 253-262.

Sierpinski, W. General topology. Toronto, University of Toronto Press, 1952.

Skinner, B. F. The science of learning and the art of teaching. Harvard educ. Rev., 24, 1954, 86-97.

Smith, D. E. History of modern mathematics. New York, McGraw-Hill Book Co., 1929.

Smith, R. R. Meaningful division. Math. Tchr., 43, 1950, 12-18.

Sokolnikoff, I. S. Advanced calculus. New York, McGraw-Hill Book Co., 1939.

Spearman, C. E. The abilities of man, their nature and measurement. London, Macmillan & Co., 1927.

Stone, M. M. Some crucial problems in mathematical instruction in the limited states. Sch. Rev. 65, 1957, 64-78.

Taylor, J. G. Experimental design, a cloak for intellectual sterility, Brit. J. Psychol., 49, 1958, 106-116.

Thiele, C. L. An incidental or an organized program of number teaching? Math. Tchr., 31, 1938, 63-67.

Thorndike, E. L. Psychology of arithmetic. New York, Macmillan, 1922.

Thorndike, E. L. The psychology of algebra. New York, Macmillan, 1923.

Thorndike, E. L. Mental discipline in high school studies. J. educ. Psychol. 15, 1924, 1-22, 83-98.

Thorndike, E. L. The measurement of intelligence. New York, Teachers College, 1927.

Thurstone, L. L. The nature of intelligence. New York, Harcourt, Brace & Co., 1927.

Thurstone, L. L. Primary mental abilities. Chicago, University of Chicago Press, 1938.

Thurstone, L. L. Factorial studies of intelligence. Chicago, University of Chicago Press, 1941.

Thurstone, L. L. Multiple factor analysis. New York, Harcourt, Brace & Co., 1947.

Thrall, R. M., Coombs, C. H., and Davis, R. L. Decision Processes, New York, Authors, Wiley, 1954.

Tuttle, L. and Satterly, J. The theory of measurements. London, Longmans, Green & Co., 1925.

Twenty-first Yearbook of the National Council of Teachers of Mathematics. The learning of mathematics, its theory and practice. Washington, D.C., National Council of Teachers of Mathematics, 1953.

Ulmer, G. Teaching geometry to cultivate reflective thinking: an experimental study with 1239 high school pupils. J. exp. Educ., 8, 1939, 18-25.

University of Illinois Committee on School Mathematics. High school mathematics, first course, teachers' edition. Urbana, Ill., University of Illinois Press, 1953.

VanDuren, W. L. (Jr.). The maneuvers in set thinking. Math. Tchr., 51, 1958, 322-335.

Veblen, O. The foundations of geometry, in J.W.A. Young, (ed.), Monographs on topics of modern mathematics. New York, Dover Publications, 1955, Chapter 1.

Vernon, D. E. The structure of human abilities. London, Meuthuen, 1950.

Vinacke, W. E. The psychology of thinking. New York, McGraw-Hill, 1952.

Wallas, G. The art of thought. London, C. A. Watts & Co., Ltd., 1926.

Walker, H. M. & Lev, J. Statistical inference. New York, Henry Holt & Co. Inc., 1953.

Washburne, C. W. The grade placement of arithmetic topics: 'a committee of seven' investigation. Report of the society's committee on arithmetic. Twenty-ninth Yearbook, part 2, National Society for the Study of Education, Bloomington, Ill., Public School Publishing Co., 1930, 641-670.

Washburne, C. W. Work of the committee of seven on grade placement of arithmetic. Child development and the curriculum. Thirty-eighth Yearbook, part 1, National Society for the Study of Education. Bloomington, Ill. Public School Publishing Co., 1939, 299-324.

Weaver, J. F. Six years of research in arithmetic instruction. Arith. Tchr. 4, 1957, 89-100.

Wertheimer, M. Productive thinking. New York, Harper, 1945.

Wesman, A. G. Expectancy tables--a way of interpreting test validity. Test Service Bull., New York, Psychological Corporation, 38, 1949.

Wheeler, R. H. The new psychology of learning. Ninth Yearbook of the National Council of Teachers of Mathematics, New York, Bureau of publications, Teachers College, Columbia University, 1935.

Wilson, K. Information theory approach to psycholinguistics. J. abnorm. soc. Psychol., 49, 1954, 35-49.

Wisely, E. An approach to problem solving. Arith. tchr., 4, 1957, 125-129.

Woods, F. S. Non-Euclidian geometry, in J. W. A. Young, (ed.), in Monographs on topics of modern mathematics. New York, Dover Publications, 1955, Chapter III.

Woody, C. Knowledge of arithmetic possessed by young children. Bull. Sch. Educ., Indiana Univer., 1930, 50-85.

Wrigley, J. The factorial nature of ability in elementary mathematics. Brit. J. educ. Psychol., 78, 1958, 61-78.

Wrightstone, J. W. Influence of research on instruction in arithmetic. Math. Tchr., 45, 1952, 182-197.

APPENDICES

APPENDIX A

THE AXIOMS OF PROJECTIVE GEOMETRY

Let us consider a class of undefined elements which we shall call 'points'. An undefined subclass of points we shall call a 'line'. It is customary to represent points by capital letters A, B, C, . . . and lines by small letters a, b, c, . . . If a point belongs to a sub-class which we have called a line, we shall say that it is 'on' that line; conversely, we shall say that the line is 'on' or 'passes through' the point. For formal statements it is convenient to use the word "on" in both senses, though "passes through" is more familiar in the latter connection. We shall say that two lines which have a point in common are 'concurrent in' or 'intersect in' that point, and that any three points of a line are collinear.

We now make the following assumptions concerning points and lines:

I. There are at least two distinct points.

II. Two distinct points A and B determine one and only one line on both A and B. This line we shall call AB (or BA), and speak of it as 'joining' A and B. It is not difficult to prove from II that if C and D are distinct points on AB, then A and B are points on CD; also, that 'two distinct lines cannot have more than one common point'.

III. If A and B are distinct points, there is at least one point distinct from A and B on the line AB.

IV. If A and B are distinct points, there is at least one point not on the line AB.

V. If A, B, C, are three non-collinear points, and D is a point on BC distinct from B and C, and E is a point on CA distinct from C and A, then there is a point F on AB such that D, E, F, are collinear.

Definition. If A, B, C are three non-collinear points, the 'plane' ABC is the class of points lying on lines joining C to the points of line AB.

VI. If A, B, C, are three non-collinear points, there

is at least one point D not on the plane ABC.

<u>Definition</u>. If A, B, C, D are four non-coplanar points, the 'three-dimensional space' ABCD is the class of points lying on lines joining D to the points of the place ABC.

VII. Any two distinct planes have a line in common.

APPENDIX B

DESCRIPTIONS OF STANDARDIZED TESTS AND QUESTIONNAIRES

1. Differential Aptitude Tests--Numerical Ability (Form A)

According to the manual, "The numerical ability items are designed to test understanding of numerical relationships and facility in handling numerical concepts. The problems are framed in the item type usually called 'arithmetical computation', rather than what is usually called 'arithmetical reasoning.'" The author's claim that the test offers a measure of numerical ability, uncontaminated by verbal elements.

Corrected split-half reliability coefficients of the order of .88 are reported in groups with standard deviations equal to 8.7. Testing time is 30 minutes.

2. Differential Aptitude Tests--Space Relations (Form A)

This test is designed to measure both the ability to visualize an object constructed from a picture or pattern as well as the ability to visualize how an object would look if related in various ways. Thirty minutes are required in actual testing time for this test.

Corrected split-half reliability coefficients and standard deviations are listed as .91 and 23.0 respectively.

3. The Dominion Advanced Group Test of Learning Capacity (Form A).

This is a 30 minute, 75 item omnibus-type intelligence test which boasts parallel form reliability coefficients of .85 (SD=10.2) at the grade 10 level and .90 (SD=12.9) at the grade 12 level. For the combined grades 10 and 12 sample (N=107), the correlation between the Dominion and Otis amounted to .89.

4. The Gordon Personal Inventory

A 15 minute inventory designed to measure four aspects of personality: Cautionsness, Original Thinking, Personal Relations, and Vigor. Percentile norms for high school are provided.

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

5. Holtzman-Brown Survey of Study Habits

The test provides a measure of study methods, motivation for studying, and certain attitudes twoard scholastic activities important in the classroom. Norms are provided for college and high school (separate norms for male and female).

Split-half reliability coefficients of .83 are quoted in the manual (SD=11.2). The test shows very low correlations with most tests of ability, but correlates reasonably high with school grades (some correlations reported as high as .66). This test would therefore seem to be useful in conjunction with a test of ability in a multiple correlation prediction of grades.

6. Kuder Preference Record--Vocational Form B

The Kuder was designed to allow a systematic approach to the problem of vocational choice in terms of expressed interest. Interests are measured in ten broad areas: 1. outdoor, 2. mechanical, 3. computationa, 4. scientific, 5. persuasive, 6. artistic, 7. literary, 8. musical, 9. social service, and 10. clerical.

Profiles may be compared with those for a wide variety of occupational groups. Reliabilities coefficients are supplied for each scale and cluster about .87.

7. Minnesota Counseling Inventory

This test, designed primarily for students in grades 9 to 12, contains many items from the Minnesota Multiphasic Personality Inventory. It requires approximately 50 minutes for administration and yields 9 scores altogether.

A validity score and a questions-not-answered score gives a preliminary screening for invalid tests. Three scores assist in identifying areas in which students may be adjusting particularly well or poorly. These are the scores on the Family Relationships (F.R.), Social Relationships (S.R.), and Emotional Stability (E.S.) scales. The remaining four scores provide information more directly related to the methods students employ in making adjustments. These scores are those of Conformity, (C.); Adjustment to Reality (R), Mood(M), and Leadership (L) scales.

Norms are provided which allow the conversion of raw scores to standard scores for grades 9-10 and 11-12, with

separate norms for boys and girls. Criterion group scores for the interpretation of each scale and for both sexes range from .56 to .93.

8. Otis Quick Scoring Mental Ability Tests (Gamma E.M.)

This omnibus-type intelligence test contains 80 items requiring 30 minutes testing time. The manual reports corrected split-half reliabilities of .92 ($S.E._{meas}=3.0$).

9. Revised Minnesota Paper Form Board Test--Series AA

The test consists of a series of two-dimensional diagrams into separate parts. For each diagram there are few figures with lines indicating the different shapes out of which they are made. From these, the subject chooses the one figure which is composed of the exact parts that are shown in the original diagram.

The test requires 20 minutes to administer and reports 'interform' reliabilities of .85 (SD not given).

10. Watson-Glaser Critical Thinking Appraisal (Form A)

A 99-item, 40 minute test designed to provide problems and situations which require the application of some of the important abilities involved in critical thinking. It contains five subtests:

Test 1. Inference. (20 items) Designed to sample ability to discriminate among degrees of truth or falsity or probability of certain inferences drawn from given facts or data.

Test 2. Recognition of Assumptions. (16 items) Designed to sample ability to recognize unstated assumptions in given assertions or propositions.

Test 3. Deduction. (25 items) Designed to sample ability to reason deductively from given premises; to recognize the relation of implication between propositions; to determine whether what seems an implication or necessary inference between one proposition and another is indeed such.

Test 4. Interpretation. (24 items) Designed to sample ability to weigh evidence and to distinguish

between unwarranted generalizations and probable inferences which, though not conclusive, or necessary, are warranted beyond reasonable doubt.

Test 5. Evaluation of Arguments. (14 items) Designed to sample ability to distinguish between arguments which are strong and important to the question at issue and those which are weak and unimportant or irrelevant.

Percentage norms are provided for high school and college students. Whole-test interform reliabilities ranging from .81 (SD=12.0) to .93 (SD=14.0) are reported. The reliability coefficients of the individual scales range from .36 to .78.

Subject Preferences and Level of Aspiration

Name: ______________________________
Grade: _____________________________

Instructions: Fill in the following questionnaire by making entries to the following directions:

Column 1: Rank the subjects listed in the order (1, 2, 3, etc.) that you enjoy them most, with your best-liked subject marked 1, etc.

Column 2: Rank the subjects according to the marks that you receive in each, with your highest subject marked 1, etc.

Column 3: Rank the subjects according to the time spent in each one, with the subject requiring most time ranked 1, etc.

Column 4: Enter beside each subject the mark that you would be satisfied with in that subject.

	Enjoy Most	Highest Marks	Spend Most Time at	Expected Mark
Science				
English				
French				
Latin				
Mathematics				
P.T.				
Shop (H.Ec.)				

Socio-Economic Data

Student Information Sheet

1. Name________________________ Age: _____Yrs._____Months
2. Present Address______________________________________
3. How many years have you lived at this address?_________
4. Address prior to this?________________________________
5. Do you live with your parents?________
6. Father's Occupation__________Mother's Occupation_______
7. No. of younger brothers and sisters_______Older________
8. Religious Denomination________________________________
9. How far do you intend to go in high school?____________
10. Would you be interested in attending university?_______
11. Could you afford to go to university?__________________
12. What kind of work do you intend to do when you complete your education?______________________________________
13. What grade did your parents reach in school?

    Father____________ Mother____________
14. What is the highest grade reached by any of your brothers or sisters?__________________

Geometry Background Data

Student Information Sheet

1. Do you find this year's Geometry to be one of your easier or harder subjects?__________

2. Did you find Algebra to be easier or harder than Geometry? __________

3. Did you find public school arithmetic to be one of your best or worst subjects?__________

4. When did you become aware that mathematics was going to be easy (or difficult) for you?__________

5. Can you offer any reason why you should find mathematics easy or difficult?__________

6. Did either of your parents (brothers or sisters) find mathematics difficult? Elaborate__________

7. Was either of your parents (brothers or sisters) very strong in mathematics? Elaborate__________

8. When you entered high school did you expect to find mathematics difficult?__________

9. Do you think that mathematics requires a special ability or talent?__________

10. Do you think that the solution of a mathematics problem requires more intelligence than the understanding of a poem?__________

11. Do you think that boys should do better than girls (of equal intelligence) in mathematics?__________

12. Which of the following reasons for studying mathematics most nearly applies to you?
    - a. the desire to obtain a high mark (or pass).
    - b. pleasure derived from solving problems.
    - c. have to study because on course of study.

    __________(be honest)

13. In what way do you think that a student will profit from a study of geometry? Explain.______________________________
____
____

14. What mark do you think would fairly represent your work in geometry this term?__________

15. What mark would you be satisfied with for this year's work?__________

16. What is the longest time that you have ever spent trying to solve a problem? Elaborate.___________________________
____

APPENDIX C

LEVEL I TESTS

Number Series--Level I

A. Step-level 1

1. 2, 5, 8, __
2. 2, 7, 12, __
3. 15, 11, 7, __
4. 22, 16, 10, __
5. 5, 13, 21, __
6. 4, 10, 16, __
7. 22, 15, 8, __
8. 17, 13, 9, __
9. 3, 9, 15, __
10. 2, 12, 22, __
11. 11, 8, 5, __
12. 3, 8, 13, __
13. 2, 6, 10, __
14. 3, 11, 19, __
15. 30, 21, 12, __
16. 22, 16, 10, __
17. 1, 8, 15, __
18. 5, 9, 13, __
19. 21, 15, 9, __
20. 3, 13, 23, __

B. Step-level 2

1. 1, 3, 7, 13, __
2. 2, 3, 7, 14, __
3. 1, 7, 12, 20, __
4. 3, 5, 9, 15, __
5. 2, 5, 10, 17, __
6. 3, 5, 10, 18, __
7. 3, 7, 13, 21, __
8. 2, 5, 9, 14, __

C. Step-level 3

1. 2, 5, 10, 18, 30, __
2. 1, 2, 4, 13, 29, __
3. 5, 6, 8, 12, 19, __
4. 3, 4, 7, 15, 31, __

D. Step-level 4

1. 2, 4, 7, 12, 21, 27, __

2. 2, 3, 7, 16, 33, 62, __

3. 1, 2, 4, 9, 20, 41, __

4. 3, 4, 7, 13, 25, 48, __

E. Step-level 5

1. 1, 2, 5, 11, 23, 47, 93, __

2. 2, 4, 8, 15, 28, 52, 98, __

3. 1, 2, 4, 8, 16, 32, 63, __

4. 2, 3, 5, 9, 19, 43, 94, __

F. Step-level 7

1. 2, 4, 7, 13, 25, 49, 97, 192, 376, __

2. 1, 2, 4, 8, 16, 32, 64, 129, 263, __

3. 1, 2, 5, 11, 22, 42, 81, 163, 339, __

4. 1, 2, 4, 9, 21, 47, 100, 206, 417, __

Permutations and Gap Filling

A. Practice Set: a b c

Permutations	Gap Filling	
I_{12}	b a c	
I_{13}	c a b	G_1
I_{23}	a c b	
$I_{23, 13}$	b c a	G_2
$I_{12, 13}$	c a b	

B. Test Set

Permutations	Gap Filling	
1. I_{24}	7. a d c b	
2. I_{12}	8. a b d c	
3. I_{14}	9. b a c d	G_1
4. I_{34}	10. a c b d	
5. I_{13}	11. d b c a	
6. I_{23}	12. c b a d	

Permutations	Gap Filling	
13. $I_{24,14}$	21. b d c a	
14. $I_{23,34}$	22. a c d b	
15. $I_{23,24}$	23. a d b c	
16. $I_{12,14}$	24. d a c b	G_2
17. $I_{13,23}$	25. c a b d	
18. $I_{34,14}$	26. c b d a	
19. $I_{12,23}$	27. b c a d	
20. $I_{13,14}$	28. d b a c	

29. $I_{12,24,23}$	35. b c d a	
30. $I_{34,14,24}$	36. c d b a	
31. $I_{12,13,14}$	37. d a b c	G_3
32. $I_{34,23,13}$	38. b d a c	
33. $I_{23,12,24}$	39. c d b a	
34. $I_{14,34,24}$	40. b c d a	

Ordering

Practice Card A: A is larger than G.
A is smaller than J.
Which is the smallest?

Practice Card B: X is smaller than Y.
W is larger than Y.
Which is the largest?

Practice Card C: N is smaller than K.
K is smaller than X.
Z is smaller than N.
Name the letters in order of size.

1. K is smaller than R.
R is smaller than L.
Which is the smallest?

2. H is larger than L.
H is smaller than O.
Which is the largest?

3. A is smaller than Z.
K is larger than Z.
Which is the smallest?

4. R is larger than J.
G is larger than R.
Which is the largest?

5. N is larger than Q.
N is smaller than W.
Which is the smallest?

6. H is smaller than N.
N. is smaller than S.
Which is the largest?

7. C is larger than M.
K is larger than C.
Which is the smallest?

8. U is smaller than V.
M is larger than V.
Which is the largest?

9. A is smaller than B.
D is larger than B.
C is smaller than A.

10. X is larger than K.
N is smaller than Z.
X is smaller than N.
K is larger than Y.

11. N is larger than B.
B is larger than R.
P is larger than H.
H is larger than N.
T is larger than P.

12. K is larger than J.
F is larger than L.
X is smaller than S.
L. is larger than K.
M is smaller than J.
X is larger than F.

13. N is smaller than A.
V is larger than Z.
L is smaller than M.
A is smaller than Z.
P is larger than F.
N is larger than P.
L is larger than V.

Topological Comprehension Test

1. Set and Subset

   (a) Name four subsets of the set (a, b, c, d).

   (b) Name three subsets of the set consisting of all the students in the first row.

2. (V) Space

   (a) Construct a (V) space consisting of five elements in which each element has a different number of neighbourhoods.

3. Limit Point

   Define neighbourhoods of the space (a, b, c, d) so that the element 'c' is the only limit point of the subset (a, c, d).

4. Derived Set

   Define neighbourhoods of the space (a, b, c, d) so that the derived set is null.

5. Topological Equivalence

   Test the following (V) spaces for topological equivalence:

V_1 (a, b)			V_2 (a, b)		
nhds. of a:	a ,	a, b	nhds. of a:	a ,	a, b
nhds. of b:	b		nhds. of b:	b ,	a, b

6. Theorem 1

   Demonstrate the 'necessary' part of the theorem, using the topologically equivalent spaces whose neighbourhoods are defined as follows:

V_1 (a, b)		V_2 (a, b)		
nhds. of a:	a	nhds. of a:	a ,	a, b
nhds. of b:	a, b	nhds. of b:	a, b	

APPENDIX D

STUDIES AT LEVEL II

Axioms and Definitions in Grade 10 Geometry

Undefined terms: point, line.

Defined terms: (angle, triangle, circle
(straight angle contains 180° (A)
(congruency

Axioms:

(1) Equality

(2) Addition

(3) Subtraction

(4) Substitution

(5) Multiplication

(6) Division

(7) Completion

(8) Coincidence

Propositions:

I. If two straight lines intersect, the vertically opposite angles are equal.

II. If two sides and the contained angle of one triangle are respectively equal to two sides and the contained angle of another triangle, then the triangles are congruent.

III. In an isosceles triangle, the angles opposite the equal sides are equal.

IV. If three sides of one triangle are respectively equal to three sides of another triangle, then the two triangles are congruent.

XIII. If a transversal cuts two straight lines making the alternate angles equal, then the lines are parallel.

XIV. If a transversal cuts two parallel straight lines, then the alternate angles are equal.

XVI. The sum of the angles of a triangle equals 180°.

XVII. If two angles and one side of a triangle are respectively equal to two angles and one side of another triangle, then the triangles are congruent.

XVIII. If two angles of a triangle are equal, the sides opposite the equal angles are equal.

XIX. If the hypotenuse and one side of a right-angled triangle are respectively equal to the hypotenuse and one side of another right-angled triangle, then the triangles are congruent.

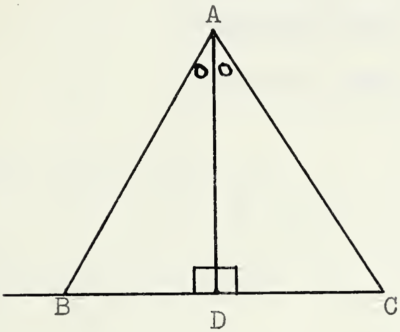
XXII. If two sides of a quadrilateral are equal and parallel, then the other two sides are equal and parallel.

XXIII. In a parallelogram, the opposite sides are equal, the opposite angles are equal, etc.

NAME________________ GIVEN: DIAGRAM TIME________

GRADE_______________ PROVE: AB=AC

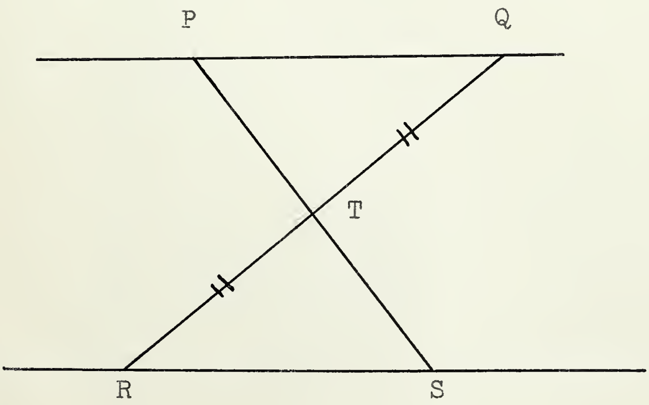
(59)



---

(60) GIVEN: PQ RS TIME________

PROVE: PT=TS



TIME________

Number Series--Level II

Practice Card A: 1, 3, 5, __.

Practice Card B: 2, 4, 8, __.

Practice Card C: 1, 2, 1, 2, 1, 2, __.

Practice Card D: 17, 13, 9, __.

A,1:	2, 5, 8, __.	B,1:	11, 8, 5, __.
A,2:	2, 1, 1, 2, 1, 1, __.	B,2:	1, 3, 3, 1, 3, 3, __.
A,3:	1, 3, 7, 13, __.	B,3:	2, 5, 10, 17, __.
A,4:	3, 6, 12, __.	B,4:	24, 12, 6, __.
A,5:	2, 7, 12, __.	B,5:	3, 8, 13, __.
A,6:	5, 10, 15, __.	B,6:	3, 6, 9, __.
A,7:	2, 5, 10, 18, 30, __.	B,7:	5, 6, 8, 12, 19, __.
A,8:	15, 11, 7, __.	B,8:	2, 6, 10, __.
A,9:	40, 20, 10, __.	B,9:	5, 10, 20, __.
A,10:	2, 3, 7, 14, __.	B,10:	3, 5, 10, 18, __.
A,11:	2, 4, 7, 12, 21, 37, __.	B,11:	1, 2, 4, 9, 20, 41, __.

PROBLEM SOLVING SHEET

Prove	Use Proposition I	II	III	IV			
$\angle = \angle$	✓	✓	✓	✓			
$\ell = \ell$		✓					
$\triangle \cong \triangle$		✓		✓			

HAVE YOU ASKED THESE QUESTIONS

1. What kind of fact am I to prove?
2. What propositions can I use? (See Table)
3. Where can this proposition be applied?
4. What facts would I need before I can use it?